华章数学译丛



62

WILEY

Financial Statistics and Mathematical Finance

Methods, Models and Applications

金融统计与数理金融

方法、模型及应用

[德] 安斯加尔·斯特兰（Ansgar Steland）著

冉启康 尤成其 刘诚霖 译



机械工业出版社
China Machine Press

华章数学译丛
62

Financial Statistics and Mathematical Finance

Methods, Models and Applications

金融统计与数理金融

方法、模型及应用

[德] 安斯加尔·斯特兰 著
(Ansgar Steland)

冉启康 尤成其 刘诚霖 译

机械工业出版社
China Machine Press

图书在版编目（CIP）数据

金融统计与数理金融：方法、模型及应用 /（德）安斯加尔·斯特兰（Ansgar Steland）著；冉启康，尤成其，刘诚霖译．—北京：机械工业出版社，2017.6
（华章数学译丛）
书名原文：Financial Statistics and Mathematical Finance：Methods，Models and Applications

ISBN 978-7-111-57301-2

I. 金… II. ①安… ②冉… ③尤… ④刘… III. ①金融统计 ②金融学－数理经济学 IV. F830

中国版本图书馆 CIP 数据核字（2017）第 160034 号

本书版权登记号：图字：01-2012-7783

本书讨论了金融中统计方法应用的方方面面以及金融应用中统计工具使用的多种途径．首先简要介绍了金融统计和数理金融的基础知识，接着阐释了经济和金融工程中统计方法的应用和重要性，最后阐述了鞅理论、随机过程、随机积分等高级论题．本书适合统计学、金融数学、计量经济学、商务管理等专业的研究生和相关领域的从业者和研究人员阅读，许多章节也适合具有微积分和概率统计基础的本科生阅读．

出版发行：机械工业出版社（北京市西城区百万庄大街 22 号 邮政编码：100037）
责任编辑：和 静　　责任校对：李秋荣
印　　刷：三河市宏图印务有限公司　　版　　次：2017 年 7 月第 1 版第 1 次印刷
开　　本：186mm×240mm 1/16　　印　　张：20.75
书　　号：ISBN 978-7-111-57301-2　　定　　价：85.00 元

凡购本书，如有缺页、倒页、脱页，由本社发行部调换
客服热线：（010）88378991 88361066　　投稿热线：（010）88379604
购书热线：（010）68326294 88379649 68995259　　读者信箱：hzjsj@hzbook.com

译 者 序

自20世纪50年代以来，随着金融学理论与金融市场和工具的不断发展，数学与统计在金融研究中的作用显得越来越重要. 反之，统计与数学理论及方法在金融学中的应用又极大地促进了金融理论的发展，数理金融学由此孕育而生. 数理金融是20世纪80年代末90年代初出现的一门新兴学科，它是以现代数学、统计及金融理论为基础，综合利用数学模型、数值计算等开发、设计金融产品，创造性地解决各种金融问题的一门学科，其核心内容是研究不确定随机环境下的投资组合的最优选择理论和资产的定价理论. 套利、最优与均衡是数理金融的基本经济思想和三大基本概念.

本书是一部近期在西方国家非常流行的著作. 它使用时间序列分析、随机过程、随机分析等理论详细地介绍了离散时间与连续时间的金融衍生产品定价理论. 作者以高超的手法对金融衍生产品定价中所需的统计与数学理论进行了全面的描述和总结，并系统地介绍了金融衍生产品的常用定价方法和最新进展.

本书内容丰富、推导严谨、案例翔实. 由于作者在内容选择、结构安排和逻辑体系设计方面的精巧构思，所以能以相对较少的篇幅，把书中所讨论的问题的经济背景以及解决这些问题的数学方法和基本思想，系统而又简明地展示给读者，且具有相当的深度. 本书是为金融市场的量化所需的一些最重要的统计及数学理论提供的一本系统而且深入的教材，它包含了非常前沿的金融统计与数理金融. 目前，为期权这种未定权益定价的数学理论——数理金融，与分析来自金融市场数据的统计方法与统计模型的理论——金融统计这两个领域在发展过程中或多或少地相互分离了，而同时覆盖这两个领域的教材非常缺乏，而此书将这两个领域有机地整合到了一起. 本书适合于硕士生、博士生，高水平的研究人员以及对数理金融或金融统计感兴趣的实践工作者阅读. 其中许多章节也适合已经学习了微积分、概率论及统计的本科生阅读.

受机械工业出版社华章公司之托，我们将此书译成中文. 全书由上海财经大学数学学院冉启康教授，研究生尤成其、刘诚霖共同翻译，王清华老师参与了校对工作。限于时间和水平，译文的不当之处在所难免，敬请本书的读者和有关领域的专家批评指正.

前　言

本书系统而且深入地介绍了金融市场的量化所需的一些最重要的数学理论，它包含了狭义的数理金融：为期权这种未定权益定价的套利理论及相关的数学理论，以及分析来自金融市场数据的统计方法与统计模型. 这两个领域在发展过程中或多或少地相互分离了，而同时覆盖这两个领域的教材又非常缺乏，这是我写作本书的主要动机. 我尝试着实现这一目的. 本书适合硕士生、博士生、研究人员以及对上述两个领域感兴趣的实践工作者阅读. 其中许多章节也适合已经学习了微积分、概率论及统计的本科生阅读. 除了少数例外，所有的结果都给出了详细的证明，尽管可能与传统的证法有所不同. 为了避免过多的概念、记号、模型和方法使得教材变得太复杂，也为了读者在首次阅读的时候就能跟随计算和推导快速地掌握本书的内容，我们尽可能地使本书的数学公式及记号初等化. 在每章的结尾部分，我们都给出了所用参考文献的评注，这有助于更好地理解本书，也为进一步的学习提供了参考.

第 1 章从介绍一些重要的概念，如期权、金融衍生品等金融工具及相关的交易策略开始. 但本章的重点不是阐述衍生品的原理和基本结果，而是引入后续章节所需的各种金融术语. 本章也给出了现金流、贴现和利率期限结构的初步介绍. 在一段给定时期上的资产收益（通常是日均收益）是金融市场中非常重要的研究主体，因为资产可根据收益重新定价，评判投资的依据也是资产的收益. 对于收益的预估、预估误差、扰动范围等量的统计学分析有着重要的经济意义，所以，本章对相关的统计估计方法作了详细的介绍. 要对投资风险进行度量必须知道相关统计估计的性质. 例如，波动率与投资收益的标准误差密切相关，而风险价值，顾名思义，需要研究风险的量化及它们的统计估计. 第 1 章以初步介绍期权定价作结尾，引入了数理金融学中最重要的一些概念，如无套利原理、风险中性定价原理，以及这些概念与概率演算（特别是等价鞅测度的存在性）的关系. 事实上，通过最基础的介绍或一些简单的实例，就可以迅速掌握以上概念和基本结论.

第 2 章讨论套利理论和单期模型的未定权益的定价. 设在 0 时刻，投资者建立了一个资产组合，需要讨论该资产组合在 1 时刻的收益状况. 在一个简单的框架内，本章对第 1 章的结果作严格的数学处理，并将结果从有限概率空间（在该空间假设下，市场仅有有限种情况发生）推广到一般概率空间中去，使之更符合实际市场. 空间分离定理表明，任意给定一个点，我们可以把它和指定的凸集分离. 分离定理被用来证明套利机会消失及等价鞅测度的存在性，因此本章给出了分离定理的详细介绍，并给出了建立在 Esscher 变换下的等价鞅测度的构造方法.

第 3 章详细地介绍了离散时间（时间序列）的随机过程，包括鞅、鞅差分序列、线性过程、ARMA 过程、GARCH 过程，以及长记忆序列. 鞅是数理金融中的一个基本概念，研究结论表明，在任意无套利的金融市场中，均存在一个概率测度，使得风险资产的贴现价格过程在该测度下是一个鞅，而任意一项未定权益，均能在该测度下得到它的风险中性定价，鞅理论对得到这些成果起到了关键性的作用. 但是本章仅限于介绍在后面各章中需

要用到的鞅性质. 对鞅过程取一阶差分，即*鞅差分序列*，白噪声是它的一种形式，通常用来代替金融随机模型甚至经济随机模型中误差项是独立同分布这种不符合现实的假定. 统计分析表明，金融收益序列可以被假定为不相关的，但通常它们不是独立的. 当然具有相关性的一些时间序列也需要纳入考虑. ARMA 时间序列模型是一类适用范围较广的参数模型，它是一般的无限维线性过程，本章介绍了 ARMA 模型的参数和自协方差函数的估计方法. 许多金融时间序列还有条件异方差性，为此引出了 GARCH 模型. 本章的最后介绍了分数阶差分和长记忆过程.

第 4 章详细介绍了离散时间的多期模型的套利理论，在本章的模型中假设交易发生在一系列有限的时刻，在每个时刻，投资者均可以利用已知的市场信息来调整资产组合. 可以应用第 3 章的离散时间的鞅理论来研究无套利金融市场上期权和其他衍生品的定价. 本章详细研究了 CRR 二叉树模型，该模型是实际应用中的标准模型之一，由它可导出著名的欧式看涨期权的 Black-Scholes 定价公式. 除此之外，本章还讨论了美式权益的定价方法，其中需要用到最优停时的高等数学理论.

第 5 章介绍连续时间的随机过程. 布朗运动是金融市场连续时间价格模型的主要随机源，为了使得内容简洁，本章仅限于讨论布朗运动的定义及最重要的一些性质. 布朗运动有一些令人惊讶的性质，比如：路径连续但处处不可微，或不存在有界变差. 本章还分别介绍了布朗运动的推广模型——分数布朗运动和 Lévy 过程. Lévy 过程保留了独立增量的性质，但是允许其增量是非正态分布的，包括其增量可能是厚尾分布并带有跳的. 与布朗运动一样，分数布朗运动也是一个高斯过程，但分数布朗运动可能有长相依的增量，即相关性减少得非常慢.

第 6 章介绍随机积分理论. 在默认读者已经掌握 Riemann 积分和 Lebesgue 积分的基础上，本章从介绍 Riemann 积分的直接推广——Riemann-Stieltjes（RS）积分开始，这种积分相对来说比较容易，它为引入 Itô积分作了铺垫. 值得一提的是 RS 积分可以用来研究许多统计问题. 可是如果没有 Itô积分，就无法继续讨论数理金融. 主要原因在于：在时间 $[t,t+\delta]$ 内，股票数为 $x(t)=x_t$ ，资产变化为 $x_t\delta P_t$ ，其中，$\delta P_t=P_{t+\delta}-P_t$. 在 n 期时间 $[i\delta,\ (i+1)\delta](i=0,\ 1,\ \cdots,\ n-1)$ 上资产的累积变化为 $\sum_{i=0}^{n-1}x(i\delta)\delta P_{i\delta}$. 现取极限 $\delta\to 0$ ，就出现了一个关于股票价格的积分 $\int x_s\mathrm{d}P_s$ ，但是当股价不是有界变差过程时，这个积分在 Stieltjes 意义下是没有意义的，这就必须要引入 Itô积分. 本章还将介绍应用中常见的 Itô过程. Itô公式表明：一个 Itô过程的光滑函数仍是 Itô过程，且这个 Itô过程可以具体表示出来. 作为一类重要的 Itô过程，遍历扩散也将在本章中介绍. 本章介绍的 Euler 数值逼近方法为离散样本遍历扩散的统计估计及统计推断奠定了基础.

第 7 章介绍衍生品定价的一个理想的数学模型，即 Black-Scholes 模型，在实务中，它仍是连续时间模型的基准. 在本章的模型中，投资者可以把资金投资于有风险的股票，也可以存入银行获得固定利息. 第 6 章介绍的 Itô积分为建立连续时间模型的套利理论提供了数学基础. 经典的 Black-Scholes 模型假定股价的波动率与时间无关，但现实中常常不是这

样. 因此，本章还简要地讨论了波动率是与时间有关但是非随机的情形. 最后，本章介绍推广的 Black-Scholes 模型，在这种模型中，它允许无风险工具的利率随时间变化且是随机的，这包含了不投资于股票的资金可以购买比如 AAA 级政府债券这一现实情形.

第 8 章介绍离散时间过程的渐近极限理论，这种离散过程用来构建所需的模型；经常用金融数据（如收益、指数、价格和风险度量等）来估计、推断及检验模型就是这样一种过程. 极限理论包括了鞅差、线性过程以及混合过程的大数定律和中心极限定理，其内容还在不断扩展. 本章还深入讨论了带随机回归量的多元线性回归、非参数密度估计、非参数回归以及自协方差和长期方差估计等内容，这些统计工具在金融数据分析中处处用到.

第 9 章讨论了一些特定的专题. Copula 函数已成为对高维分布建模的一个重要工具，在应用于信用及违约相关的金融工具的定价时，它是强有力的，但同时也是危险的. 事实上，在 2008 年的金融危机中它们扮演了一个不幸的角色，在那个时候，这个简单的定价模型被大规模地应用于信用违约债务的定价. 对导致这场危机的原因进行回顾，它揭示了金融市场固有的复杂性以及对复杂数学模型的需求. 本章将详细地介绍局部多项式估计，因为它在许多金融问题中都有重要的应用，比如风险中性密度条件波动率的估计或有离散观测值的扩散过程的估计. 渐近正态性是建立在下列强大的还原原理基础上的：扰动项 $\{\varepsilon_t\}$ 以及回归量的驱动过程的（联合）平滑中心极限定理蕴含了局部线性估计量的渐近正态性. 变点（结构性改变）检验和监测已成为当前理论和应用研究的热点，本章的最后简要地介绍了变点分析和监测，其中主要介绍单整阶数改变的监测.

本书附设一个专用网站：http://fsmf. stochastik. rwth-aachen. de.

致谢

我要衷心地感谢我的学生和助理 Martin Euskirchen、Wolfgant Herff 博士和 Annabel Prause 硕士，是他们仔细地阅读了此书的初稿，Martin Euskirchen 还帮助输入了部分章节. 我还要特别感谢 Mohammed Abujarad 博士，他校对了本书的最后版本，发现了一些写作错误和许多的打印错误，并为最终的成书提出了几个建议. 最后，我要感谢我的家人，特别是我的孩子们对我的支持.

Ansgar Steland
于德国 RWTH Aachen 大学

目　录

第 1 章　金融微积分基础

1.1 几个引例

▶例 1.1.1　设有一个向职工收取保费的养老基金，该基金计划把部分资金投资于交易所的股票，而不是国库券．国库券的固定利率可以事先获知，而股票收益却是不确定的．它可能大大超出债券利率，也可能因股价下跌而使基金遭受损失．对养老基金来说，了解该投资带来的预期收益和相应的投资风险是至关重要的，这有助于确定股票的投资额度．实际市场中，投资者会投资于一系列的风险资产组合．其中引出的问题是：资产组合之间的相互关系是什么？现代金融学用服从特定分布的随机变量来描述收益．因此，必须从数学和统计学的角度考虑收益满足的特征，从经济学的角度考虑收益的期望、收益的风险，并给出精确的度量．进一步，在给定历史数据的时间序列后，需考虑如何得到序列的预测值． ◀

▶例 1.1.2　为了减少股票投资可能带来的损失，养老基金可以和银行签订合约，当合约敲定的止损价 L 大于股价 S 时，基金可以对该合约行权，要求银行支付价差，这种合约被称为期权．如何为该期权作公平定价，银行为对冲期权空头应如何确定投资策略？ ◀

▶例 1.1.3　设某钢铁厂为年产量为 10 000 辆的汽车制造厂提供必需的钢材，钢铁厂年内开始生产，需要耗用大量石油．为了计算成本，钢铁厂希望能事先锁定油价，不妨设为 K．一个可行的方法是，钢铁厂可以签订一项合约，在合约的交割期，如果实际油价高于 K 值，那么钢铁厂可以要求对方支付价差．这类合约被称为看涨期权．同样地引出一个问题：合约的公平定价如何确定？钢铁厂锁定油价的另一个可行方法是远期或期货合约． ◀

▶例 1.1.4　更具体地讨论例 1.1.3，设钢铁厂需要一桶石油，在当前时刻 $t=0$，石油现价 $S_0=100$．为了锁定油价，钢铁厂购买了一份买方看涨期权，执行价 $K=100$．不妨设固定利率为 1%，同时设一年后，即在时刻 $t=1$ 时的油价 S_1 满足如下的两点分布，

$$P(S_1=110)=0.6,\quad P(S_1=90)=0.4.$$

如果 $S_1=110$，钢铁厂对期权行权可获利 $G=10$，否则期权收益为 0．因此，期望收益为：

$$E(G)=10\times 0.6=6.$$

由于期权购买者具有非负收益，并且能以正的概率获得实际收益，因此，钢铁厂需要给出售期权的银行支付溢价．则银行应当以预期收益 6 作价出售该期权吗？令人惊讶的是，答案是否定的．事实上，期权的中间交易商可以一个更低的价格(例如 $x=5.45$)出售期权而不遭受损失．设交易商在 $t=0$ 时出售一份期权并以现价 50 购买 0.5 单位石油．资金流为：出售期权进账 x，购买石油负债 50．在 $t=0$ 时，投资组合包括货币市场的 $x-50$ 和产品市场的 0.5 单位石油．假定 $x<50$，那么，在 $t=1$ 时，交易商需要支付 $1.01\times|x-50|$ 给银行．分别考虑油价上涨和下跌两种情况．若油价上涨，石油升值为 $0.1\times 110=55$，此时钢铁厂行权，交易商从钢铁厂获得 100 元．则溢价 x 的合理取值是使得净支出等于购买的石

油的价格，即满足方程

$$100 + 1.01 \times (x - 50) = 55,$$

解得 $x=5.445\,545\approx 5.45$. 现在考虑油价下跌至 90 的情况，此时钢铁厂不行权转而从现货市场购买石油. 交易商仍然需要偿还借款，并低价出售石油而带来损失 5. 则溢价 x 必须保证净损益为 0，即满足方程

$$0.5 \times 90 + 1.01 \times (x - 50) = 0.$$

同样得到 $x=5.445\,545$，注意到两个方程得到了一致的非随机解 x. ◀

1.2 现金流、利率、价格和收益

首先引入一些基本概念和公式. 任一开始于时刻 $t=t_0$，终止于 T 的投资，都对应一个从数学角度描述该项投资的现金流的银行账户. 标准记法如下：记发生支付的时刻分别为 $0=t_0<t_1<\cdots<t_n=T$，对应的支付为 X_1，…，X_T. 同时约定：若支付 $X_i>0$，表示账户有进账；若支付 $X_i<0$，表示有支出.

2

从经济学角度讲，现在和将来同一数额的支付存在着巨大的差异. 为了比较该差异，考虑利息因素，或对支付贴现，将支付累计至时刻 t^*. 如果所有支付均被贴现至 $t^*=t_0$ 时刻，那么该贴现量被称为**现值**；当然，也可以选择将支付累计至时刻 $t^*=T$ 比较.

在实际中，必须确定如何分割时段，比较时间点 t_1 和 t_2 之间的经济学差异，最常用的做法是以年为基数分段. 假设日期以日-月-年的习惯给出，即 $t=(d, m, y)$. 记日期 t_1 和 t_2 之间的差异为 $\tau(t_1, t_2)$，计算 $\tau(t_1, t_2)$ 常用的基准如下：

(i) Actual/365：每年计 365 天，以实际天数计算；

(ii) Actual/360：每年计 360 天，以实际天数计算；

(iii) 30/360：每年计 360 天，每月计 30 天.

在后面的章节中，假定所有时间均服从以上基准.

设固定年利率为 r，利息支付不采取复利形式，则日期 t_1，…，t_n 对应的支付 X_1，…，X_n 在 $t=T$ 时的价值有如下表达式

$$V_T = \sum_{i=0}^{n} X_i(1+\tau(t_i,1)r).$$

而 $t=0$ 时的现值可由以下公式计算：

$$V_0 = \sum_{i=0}^{n} X_i D(0,t_i), \quad 其中 \quad D(0,t_i) = \frac{1+\tau(t_i,T)r}{1+rT}.$$

这里 $D(0, t_i)$ 表示时刻 t_i 对应的支付 X_i 的贴现因子.

通常，利息是在年内某个固定时段支付的，比如说，按季或按月支付. 把一年分成 m 个时段，同时在每一时段赋予利率 r/m，则一单位的投资经过 k 个时段后增值为：

$$1 + \frac{r}{m}k.$$

若考虑复利，则价值变为

$$(1+r/m)^k.$$

当 $k=m\to\infty$ 时，该离散利率序列趋于连续复利，即

$$\lim_{m\to\infty}(1+r/m)^m = \mathrm{e}^r.$$

3

因此，若一项投资持续 $t\in(0,\infty)$ 年，经历了 tm 个时段，那么，其累计贴现因子为

$$\lim_{m\to\infty}(1+r/m)^{mt} = \mathrm{e}^{rt}.$$

现考虑利率 $r=r(t)$ 是时间 t 的函数，且满足 $r(t)>0$，$t>0$，账户余额 $S_0(t)$ 连续增长的情形. 有两种方法可以描述这些量的关系：一种是利用带 $S_0(t)$ 的模型；一种是利用 $r(t)$. 首先假定 $S_0(t)$ 已经给出，则时段 $[t, t+h]$ 内的年增长率为：

$$\frac{1}{h}\frac{S_0(t+h)-S_0(t)}{S_0(t)}.$$

定义 1.2.1 假设账户余额 $S_0(t)$ 是可微函数，则

$$r(t) = \lim_{h\downarrow 0}\frac{1}{h}\frac{S_0(t+h)-S_0(t)}{S_0(t)}$$

有定义，称为**即时利率**或者**瞬时利率**.

由上式知：

$$r(t) = \frac{S_0'(t)}{S_0(t)} \quad\Leftrightarrow\quad S_0'(t) = r(t)S_0(t).$$

对应的微分形式为：

$$\mathrm{d}B(t) = r(t)B(t)\mathrm{d}t.$$

众所周知，该常微分方程的通解为 $S_0(t) = C\exp\left(\int_0^t r(s)\mathrm{d}s\right)$，$C\in\mathbf{R}$. 在本例中，对应于初值 $S_0(0)=1$ 的特解为

$$S_0(t) = \exp\left(\int_0^t r(s)\mathrm{d}s\right). \tag{1.1}$$

若对任意的 t，均有 $r(t)=r$，则得到 $S_0(t)=\mathrm{e}^{rt}$，与前面的结果一致.

在金融模型中，一般采用的是瞬时利率，因此常用式(1.1)来表示账户余额.

定义 1.2.2(银行账户) 一个初始资本为 1、采取即时利率 $r(t)$ 和连续复利计算的银行账户余额可表示为

$$S_0(t) = \exp\left(\int_0^t r(s)\mathrm{d}s\right),\quad t\geqslant 0.$$

若初始投资资金为 x，则在时刻 t 的资金为 $xS_0(t)$. 反过来，为了在时刻 T 得到一单位的货币，需要在 $t=0$ 时注入 $x=1/S_0(T)$ 单位的资金，初始资金 $x=1/S_0(T)$ 在任意时刻 $t\in[0, T]$ 的累积值为：

4

$$xS_0(t) = \frac{S_0(t)}{S_0(T)}.$$

上式表明，T 时刻的一单位支付在 $t=0$ 时的价值为 $S_0(t)/S_0(T)$.

定义 1.2.3 时刻 t 到 T 之间的**折现因子** D 是使得 t 时刻的 D 单位资金在无风险的投资回报下等于 T 时刻的一单位资金，它的表达式为：

$$D(t,T) = \frac{S_0(t)}{S_0(T)} = \exp\left(-\int_t^T r(s)\mathrm{d}s\right).$$

1.2.1 债券和利率期限结构

前面讨论的基础知识可用于债券定价和理解利率期限结构理论.

零息券是一种在**到期日**按**面值**(记为 X)支付一定数额货币的债券，零息券也称为**折扣券**. 本节和以后章节中，均假定债券是由政府发行，因此可以忽略其违约风险，同时假设时间以年为计量单位，年利率为 r. 由前面的讨论知，支付 X 的现值为

$$P_n(X)=\frac{X}{(1+r)^n}.$$

注意到该公式决定了债券定价和利率之间的一一对应关系. 这里的利率 r 同时是到期日 n 的**贴现率**或者**即期利率**，所谓即期利率是指合约中约定的当期利率.

现在考虑带息券：设在时刻 t_1，…，t_k，利息支付分别为 C_1，…，C_k，在到期日 T 支付现值为 X. 这一系列支付的现值等价于一串面值分别为 C_1，…，C_k，X，到期日分别为 t_1，…，t_k，T 的零息债券之和，因此，其价格由下列**债券定价公式**给出

$$P(t)=\sum_{i=1}^{k}C_iP(t,t_i)+XP(t,T),$$

如果用 $\tau_j=t_j-t$ 表示从时间点 t 到第 j 个支付的时差(即到期期限)，则上式等价于

$$P(t)=\sum_{i=1}^{k}C_iP(t,t+\tau_i)+XP(t,T).$$

由此可知，债券在时刻 t 的价格由曲线 $\tau\mapsto P(t,\ t+\tau)$决定，该曲线给出了时差为 τ，支付 1 个单位的零息债券在时刻 t 的价格. 这种方法称为**利率期限结构**.

5

也可采取第二种方法来讨论利率期限结构. 设零息券到期日为 $t+m$，面值为 1，记 $P(t,\ t+m)$为该债券在时刻 t 的价格，同时给定 m 年后到期债券在时刻 t 的即期利率 $r(t,\ t+m)$，则债券价格为

$$P(t,t+m)=\frac{1}{(1+r(t,t+m))^m}.$$

如果利息在期限内采取连续复利形式以等距 n 时段支付，则有

$$P(t,t+m)=\frac{1}{(1+r(t,t+m)/n)^{nm}}$$

当 n 变化时，上式收敛于连续复利公式. 令 $T=t+m$，则有

$$P(t,t+m)=\mathrm{e}^{-r(t,t+m)m}\quad\Leftrightarrow\quad P(t,T)=\mathrm{e}^{-r(t,T)(T-t)}.$$

连续复利率 $r(t,\ T)$也称为**到期收益率**，而函数 $t\mapsto r(t,\ T)$称为**收益率曲线**.

最后，还可以通过时刻 t 的即时远期利率来得到利率期限结构，到期日为 T 的即时远期利率定义为：

$$f(t,T)=\frac{-\dfrac{\partial}{\partial T}P(t,T)}{P(t,T)}=-\frac{\partial}{\partial T}\log P(t,T).$$

这里假定债券价格 $P(t,\ T)$关于到期期限可微，则得到

$$P(t,T)=\exp\left(-\int_0^{\tau}f(t,t+s)\mathrm{d}s\right),\quad r(t,t+\tau)=-\frac{1}{\tau}\int_0^{\tau}f(t,t+s)\mathrm{d}s.$$

1.2.2 资产收益

对于收益固定的债券，如国库券，因为利率已知，其投资价值可以事前确定．但对于股票等风险资产来说，它们的收益率是通过市场价格来计算的．

令 S_t 为时刻 t 的股价，因为股价通常是在某些等距时刻报出，一般可以将时间赋值为离散的自然数集 $\mathbf{N}$．如果投资者在时刻 $t-1$ 至 t 之间持有一份股票，则股价变化记为

$$S_t = S_{t-1}(1+R_t),$$

其中，

$$R_t = \frac{S_t - S_{t-1}}{S_{t-1}} = \frac{S_t}{S_{t-1}} - 1$$

6

称为净收益，而

$$1+R_t = \frac{S_t}{S_{t-1}}$$

称为总收益．资产收益是如何随时间积累的呢？设投资者在时刻 s 至 $t=s+k$ 的 k 个时段中持有一份股票，s，t，$k\in\mathbf{N}$（更一般地，s，t，$k\in[0,\infty)$）．定义 k-时段的收益为

$$R_t(k) = \frac{S_t - S_s}{S_s} = \frac{S_t}{S_s} - 1.$$

易知单期收益 R_{s+1}，…，R_t 和 k-时段收益之间具有如下关系：

$$1+R_t(k) = \frac{S_t}{S_s} = \prod_{i=s+1}^{t} \frac{S_i}{S_{i-1}} = \prod_{i=s+1}^{t}(1+R_i).$$

当资产持有期为 k 年，则年均收益（有效收益）为下列几何平均

$$R_{t,k} = \left[\prod_{i=0}^{k-1}(1+R_{t+i})\right]^{1/k} - 1.$$

一项具有固定的年收益 $R_{t,k}$ 的投资将会产生同样的累计收益．上式等价于

$$R_{t,k} = \exp\left[\frac{1}{k}\sum_{i=0}^{k-1}\log(1+R_{t+i})\right] - 1. \tag{1.2}$$

总收益的自然对数形式

$$r_t = \log(1+R_t) = \log\frac{S_t}{S_{t-1}}$$

称为**对数收益**．由式（1.2）可以得到从 s 到 $t=s+k$ 的 k-时段对数收益为

$$r_t(k) = \log(1+R_t(k)) = \sum_{i=s+1}^{t}\log(1+R_i) = \sum_{i=s+1}^{t} r_i.$$

和收益 R_t 相比较，对数收益具有对时间可加性的良好性质．

有了以上定义，就可以得到资产价格的下列基本乘法分解公式

$$S_t = S_0\prod_{i=1}^{t}(1+R_i) = S_0\prod_{i=1}^{t}\exp(r_i).$$

7

1.2.3 资产价格基本模型

当股票在交易所上市时(上市前认为不存在报价)，记 S_0，S_1，…为它的一系列报价，报价通常是它的市场收盘价. 其中 $S_0>0$ 表示初始报价，一般可以认为是个常数. 为了避免初始价格带来的影响，也可以事先令 $S_0=0$.

求股价随机模型的第一种方法是假设股价变化量满足

$$\Delta+u_n, \quad n=1,2,\cdots$$

其中，$\Delta\in\mathbf{R}$ 是非随机常量，$u_n(n\in\mathbf{N})$是一列独立同分布的随机变量，其分布函数为 F，并满足：

$$E(u_n)=0, \quad \mathrm{Var}(u_n)=\sigma^2\in(0,\infty), \quad \forall n\in\mathbf{N}.$$

在本书中，称 u_n 为**新息值**. 新息序列通常记为$\{u_n, n\in\mathbf{N}_0\}$，当参数 n 明确时，可简记为$\{u_n\}$. 由以上股价模型可以得到股价过程为

$$S_t=S_0+\sum_{i=1}^{t}(\Delta+u_i)=S_0+t\Delta+\sum_{i=1}^{t}u_i, \quad t=0,1,\cdots$$

其中 $u_0=0$，且约定，任意序列$\{a_n\}$有 $\sum_{i=1}^{0}a_i=0$. 称 S_t 为**随机游动**；若 $\Delta\neq0$，则称 S_t 为**带漂移项的随机游动**. 显然

$$E(S_t)=S_0+\Delta t$$
$$\mathrm{Var}(S_t)=t\sigma^2.$$

该资产价格模型主要来源于 Bachelier(1990)的工作.

另一种股价模型是建立在对数收益基础上的. 记

$$R_i := \log(S_i/S_{i-1}), \quad i\geqslant1.$$

那么

$$S_t=S_0\prod_{i=1}^{t}S_i/S_{i-1}=S_0\prod_{i=1}^{t}\exp(R_i).$$

则相应的对数价格过程为

$$\log S_t=\log S_0+\sum_{i=1}^{t}R_i, \quad t=0,1,\cdots,$$

同样，它也是一个随机游动.

对于对数收益$\{R_n\}$，经典假设认为它服从如下的正态分布：

$$R_i\overset{\text{i.i.d.}}{\sim}N(\mu,\sigma^2)$$

8 其中，$\mu\in\mathbf{R}$ 且 $\sigma^2>0$. 对数价格也同样服从正态分布：

$$\log(S_t)=\log(S_0)+\sum_{i=1}^{t}R_i\sim N(\log(S_0)+t\mu,t\sigma^2).$$

因此，S_t 服从对数正态分布. 下面对对数正态分布作简要介绍.

若随机变量 X 满足 $Y=\log(X)\sim N(\mu, \sigma^2)$，则称 X 服从参数为 $\mu\in\mathbf{R}$(称为**漂移率**)、$\sigma>0$(称为**波动率**)的**对数正态分布**，X 取值于区间$(0, \infty)$，且有

$$P(\log X \leqslant y) = \frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{y} \mathrm{e}^{-(t-\mu)^2/2\sigma^2}\,\mathrm{d}t, \quad y \in (0,\infty).$$

作变量代换 $\mu=\mathrm{e}^t$ 可得

$$P(X \leqslant \mathrm{e}^y) = P(\log X \leqslant y) = \int_{-\infty}^{\mathrm{e}^y} \frac{1}{\sqrt{2\pi}\sigma u}\mathrm{e}^{-(\log u-\mu)^2/2\sigma^2}\,\mathrm{d}u.$$

取 $y=\log x$，则得到 X 的密度函数 $f(x)$的表达式

$$f(x) = \frac{1}{\sqrt{2\pi}x\sigma}\mathrm{e}^{-(\log x-\mu)^2/2\sigma^2}\mathbf{1}(x>0), \quad x \in \mathbf{R}. \tag{1.3}$$

易知 X 的期望和方差分别为

$$E(X) = \mathrm{e}^{\mu+\sigma^2/2}, \quad \mathrm{Var}(X) = \mathrm{e}^{2\mu+\sigma^2}(\mathrm{e}^{\sigma^2}-1).$$

为了刻画比标准正态密度函数更为“矮胖”的分布，一般使用的是**自由度为 n 的 t 分布** $t(n)$，它的密度函数为

$$f(x) = \frac{1}{n\pi}\frac{\Gamma((n+1)/2)}{\Gamma(n/2)}\left(1+\frac{x^2}{n}\right)^{-\frac{n+1}{2}}, \quad x \in \mathbf{R},$$

其中，参数 n 为自然数. 由对称性知，t 分布的期望为 0，方差为 $n/(n-2)$.

这里引出几个问题：以上模型哪个是正确的或者更接近实际？收益、对数收益能服从正态分布吗？资产收益是对称分布的吗？如何估计分布的参数，如 μ，σ^2 和偏度？收益独立的假设符合实际情况吗？价格过程完全服从随机游动吗？经济状况变化将如何影响收益分布？可以发现并检验这些影响吗？如何模拟 m 个不同证券的收益之间的随机关系？

已有研究表明，某些金融变量是比正态分布尾厚的分布.

9

若随机变量 X 有**吸收域**，则称 X 是**平稳的**，或称服从**平稳分布**. X 有吸收域是指存在独立同分布的随机变量序列$\{\xi_n\}$和序列$\{\sigma_n\}\subset(0, \infty)$，$\{\mu_n\}\subset\mathbf{R}$，使得当 $n\to\infty$时，有

$$\frac{1}{\sigma_n}\sum_{i=1}^{n}\xi_i + \mu_n \xrightarrow{d} X,$$

由经典的中心极限定理知，$X\sim N(\mu, \sigma^2)$是平稳的. 由莱维-辛钦(Lévy-Khintchine)公式知，随机变量 X 的特征函数

$$\varphi(\theta) = E(\mathrm{e}^{\mathrm{i}\theta X}), \quad \theta \in \mathbf{R},$$

具有如下的表达式：

$$\varphi(\theta) = \begin{cases} \exp\left\{\mathrm{i}\mu\theta - \sigma^\alpha|\theta|^\alpha\left(1-\mathrm{i}\beta(\mathrm{sgn}(\theta))\tan\frac{\pi\alpha}{2}\right)\right\}, & \alpha \neq 1, \\ \exp\left\{\mathrm{i}\mu\theta - \sigma|\theta|\left(1+\mathrm{i}\beta\frac{2}{\pi}(\mathrm{sgn}(\theta))\log|\theta|\right)\right\}, & \alpha = 1, \end{cases}$$

其中，$0<\alpha\leqslant 2$ 称为**平稳指数**(或**特征指数**)，$-1<\beta<1$ 称为**偏度参数**，$\sigma>0$ 为**标度参数**，$\mu\in\mathbf{R}$ 为**位置参数**. 当 $\alpha=2$ 时，$\varphi(\theta)=\exp(\mathrm{i}\mu\theta-\sigma^2\theta^2)$，则为正态分布 $N(\mu, \sigma^2)$. 一个标准正态分布的尾部以指数形式快速衰减，即

$$P(|X|>x) \sim \sqrt{\frac{2}{\pi}}\frac{\mathrm{e}^{-x^2/2}}{x}, \quad x\to\infty \quad (x\sim N(0,1)).$$

相比之下，特征指数为 $0<\alpha<2$ 的平稳随机变量 X 尾部以 $x^{-\alpha}$ 速度衰减，这是因为

$$\lim_{x\to\infty} x^{\alpha}P(X>x)=C_{\alpha}\frac{1+\beta}{2}\sigma^{\alpha},\tag{1.4}$$

$$\lim_{x\to\infty} x^{\alpha}P(X<-x)=C_{\alpha}\frac{1-\beta}{2}\sigma^{\alpha},\tag{1.5}$$

其中 $C_{\alpha}=\left(\int_0^{\infty}x^{-\alpha}\sin(x)\mathrm{d}x\right)^{-1}$.

平稳分布是无穷可分分布的特殊情况. 称随机变量(或随机向量) X 和它的分布是**无穷可分的**，如果对任意的 $n\in\mathbf{N}$，存在独立同分布的随机变量 X_{n1}，X_{n2}，…，X_{nn}，使得

$$X\stackrel{d}{=}X_{n1}+\cdots+X_{nn}.$$

即无穷可分分布可以看作一列独立同分布随机变量和 $\sum_{k=1}^{n}X_{nk}$ 的分布的极限. 设 X 是 d 维随机向量，其特征函数为 $\varphi(\theta)=E(\exp(\mathrm{i}\theta'X))$，$\theta\in\mathbf{R}^d$，由莱维-辛钦公式知：

10

$$\varphi(\theta)=\exp\left\{\mathrm{i}\theta'b-\frac{1}{2}\theta'C\theta+\int_{\mathbf{R}^d}(\mathrm{e}^{\mathrm{i}\theta'x}-1-\mathrm{i}\theta'h(x))\mathrm{d}\nu(x)\right\},\tag{1.6}$$

其中，

$$h(x)=x\mathbf{1}(|x|\leqslant 1),\quad x\in\mathbf{R}^d,$$

是**截断函数**，$b\in\mathbf{R}^d$，C 为一个对称非负定 $d\times d$ 矩阵，ν 是 **Lévy 测度**，即 ν 是定义在 Borel 集 $\mathbf{R}^d$ 上的满足 $\nu(\{0\})=0$ 以及

$$\int_{\mathbf{R}^d}(|x|^2\wedge 1)\mathrm{d}\nu(x)<\infty$$

的正测度. 因此，$\varphi(\theta)$ 可由**三元组** $(b,\ C,\ \nu)$ 来刻画.

很明显，正态分布 $N(\mu,\ \sigma^2)$ 对应的三元组是 $(b,\ C,\ \nu)=(\mu,\ \sigma^2,\ 0)$. 对于强度为 λ 的泊松分布，其特征函数为

$$\varphi(\theta)=\exp(\lambda(\mathrm{e}^{\mathrm{i}\theta}-1)),$$

该函数是通过令 $b=\lambda$，$C=0$，ν 为**单点测度**(令 ν 在单个点 1 处的量为 λ)得到的.

1.3 收益的统计分析初步

前面已经介绍了价格过程可以通过收益随机变量 R_t 来描述. 为了简化讨论，假定 R_1，…，R_T 独立同分布，同时用 R 表示统一的收益，即 $R\stackrel{d}{=}R_1$，意即对任意事件 A，$P(R\in A)=P(R_1\in A)$ 成立.

在集中讨论收益之前，先简要地回顾一下任意随机变量 X 的概率性质. 不失一般性，设随机变量的分布由它的**分布函数**(d. f.)

$$F_X(x)=P(X\leqslant x),\quad x\in\mathbf{R}$$

唯一确定. 若 $f:\mathbf{R}\to\mathbf{R}$ 是满足 $\int f(x)\mathrm{d}x=1$ 的非负密度函数，则分布函数 $F(x)$ 为

$$F(x)=\int_{-\infty}^{x}f(t)\mathrm{d}t,\quad x\in\mathbf{R}.$$

具有连续密度函数 f 的随机变量 X 称为**连续型随机变量**. 通常假设收益是连续随机变量.

X 的一阶矩定义为 $\mu=E(X)$，当 X 是连续型随机变量时，一阶矩可由下式计算

$$\mu = E(X) = \int_{-\infty}^{\infty} x f(x) \mathrm{d}x.$$

11

一阶矩 $E(X)$ 又称为 X 的**均值**或**期望**. 若 X 以概率 p_1，p_2，…在离散集 $\{x_1, x_2, \cdots\}$ 上取值，即满足

$$P(X = x_i) = p_i, \quad i = 1,2,\cdots,$$

则称 X 为**离散型随机变量**，它的期望计算公式为

$$E(X) = \sum_{i=1}^{\infty} x_i p_i.$$

更一般地，可以定义 X 的 **k 阶矩** $E(X^k)$ 和 **k 阶绝对矩** $E|X|^k$. 若已知高阶矩存在，则可以估计出 X 绝对值大于某数的概率. 实际上，根据马尔可夫不等式，有

$$P(|X| > c) \leqslant \frac{E|X|^k}{c^k}.$$

即，如果高阶矩存在，那么 X 绝对值大于 c 的概率随 c 的增加而加速衰减. 将该不等式与平稳分布族的式(1.4)和式(1.5)相比较是有意义的. 若出现公司破产等负面效应，收益将出现极端离群值，**尾概率** $P(X<-c)$ 和 $P(X>c)(c>0)$ 将变得非常重要，而高阶矩则给出了该概率的取值范围.

给定随机变量的一个大小为 T 的随机样本 X_1，…，X_T，定义样本的**经验分布函数**为：

$$F_T(x) = \frac{1}{T}\sum_{t=1}^{T} \mathbf{1}(X_t \leqslant x), \quad x \in \mathbf{R}.$$

注意到 $F_T(x)$ 是随机变量小于等于 x 的观测值个数占总观测值个数的百分比.

对分布函数 F，令

$$F^{-1}(y) = \inf\{x : F(x) \geqslant y\}$$

称该左连续的反函数为**分位数函数**. 将该定义应用于经验分布函数，那么得到**样本分位数**为：

$$F_T^{-1}(p) = \inf\{x : F_T(x) \geqslant p\} = X_{(\lceil np \rceil)}, \quad p \in (0,1).$$

对固定的 p，函数值 $F_T^{-1}(p)$ 称为**样本 p 分位数**或**经验 p 分位数**. 记 $X_{(1)} \leqslant \cdots \leqslant X_{(T)}$ 为**次序统计量**，$\lceil x \rceil$ 表示大于等于 x 的最小整数，注意到 $X_{(\lceil np \rceil)} = X_{(\lfloor np \rfloor + 1)}$，其中 $\lfloor x \rfloor$ 是向下取整函数，即为小于等于 x 的最大整数. 分位数在刻画分布特征时起着重要的作用. 样本的 0.5 分位数又称**中位数**，记为 x_{med}；同样地，可以定义 0.25 分位数和 0.75 分位数

$$Q_1 = F_T^{-1}(0.25), \quad Q_3 = F_T^{-1}(0.75),$$

并称它们为**四分位数**，可以在坐标图中画出四分位点和中位数点. 这三个统计量外加最大值、最小值点可以刻画出数据集的范围. 即，基于以下 **5 个主要数据点** $x_{\min}$，Q_1，x_{med}，Q_3，$x_{\max}$，可以得到数据集的大致信息. **箱线图**以箱子和直线的形式方便地刻画了数据集，它的作法如下：画一水平数轴，在其上标出 $x_{\min}$，Q_1，x_{med}，Q_3，$x_{\max}$ 五个点，箱子以数轴为对称轴，中位数点位于 Q_1 和 Q_3 之间，并在中位数点有一分隔竖线，再从 Q_1 引一条直线至

12

$x_{\min}$，从 Q_3 引一条直线至 $x_{\max}$. 有时也把($x_{\min}$, $x_{\max}$)替换为分位点($F_T^{-1}(p)$, $F_T^{-1}(1-p)$)，p 一般取 0.01，0.05 或 0.1.

样本分位数在一般条件下是渐近正态的. 令 $p\in(0, 1)$，并记 $x_p=F^{-1}(p)$为理论 p 分位点. 若 F 在 x_p 的邻域内有正的概率密度，则当 $T\to\infty$时有：

$$\sqrt{T}(F_T^{-1}(p)-x_p)\xrightarrow{d}N(0,p(1-p)/f(x_p)^2). \tag{1.7}$$

问题出现了，渐近方差依赖于未知概率密度，它只能通过估计量 $\hat{f}_T$ 进行估计(在 1.3.4 节中将详细讨论这个问题)，希望该估计量在相当弱的正则性条件(对密度函数 f 的形状没有太多限制的条件)下具有良好的数学性质，这在分析金融数据(比如收益)时尤其重要. 基于大样本的结果式(1.7)在换成一个一致估计量后仍然成立，所以，当 $T\to\infty$时，可以直接构造出 x_p 的置信区间

$$\left[F_T^{-1}(p)-z_{1-\alpha/2}\frac{\sqrt{p(1-p)}}{\hat{f}_T(x_p)},\quad F_T^{-1}(p)+z_{1-\alpha/2}\frac{\sqrt{p(1-p)}}{\hat{f}_T(x_p)}\right],$$

当 $T\to\infty$时，该区间对应的置信概率为 $1-\alpha$，其中 $z_{1-\alpha/2}$表示标准正态分布的$(1-\alpha/2)$分位点. 在下节中将更加详细地介绍该置信区间的推导.

1.3.1 位测量

位点的度量一般以矩或分位数的形式表示，随机变量的期望是应用最多的位测量.

回到收益分析的问题，设收益分布是未知的，则期望收益 $\mu=E(R)$也是未知的. 最好的方法是利用统计估计量，即由数据 R_1，…，R_T 导出的某个函数，得到 μ 的一个好的估计. 求期望估计值的标准方法是使用所谓**经验概率测度**(即对 R_1，…，R_T 赋予同等的权重 $1/T$)下的均值去替换实际分布下的均值，因此，离散型随机变量的数学期望可认为是简单的**算术平均**：

$$\overline{R}=\overline{R}_T=\frac{1}{T}\sum_{t=1}^{T}R_t.$$

容易验证 $E(\overline{R}_T)=E(R_1)=\mu$，且不论 μ 取何值，等式恒成立. 在统计学中，满足此性质的估计量称为**无偏估计量**. 对所有可能的样本点 ω，若相应的收益为 $r=R(\omega)$，以样本点概率为权重，关于收益取加权平均，则得到与估计量相同的值 μ.

13

设有 T 个对数收益 R_1，…，R_T，现在需要检验它们的均值 $\mu=E(R)$是否等于某个特定数值 μ_0，相应的双边统计检验问题可表述为：

$$H_0:\mu=\mu_0 \quad 对应 \quad H_1:\mu\neq\mu_0.$$

假定收益是独立同分布(i. i. d.)的，且服从正态法则，则可用 t 检验方法，该检验统计量为

$$Z=\sqrt{T}\frac{\overline{R}_T-\mu_0}{S_T}, \tag{1.8}$$

其中

$$S_T=\sqrt{\frac{1}{T-1}\sum_{t=1}^{T}(R_t-\overline{R}_T)^2};$$

统计量 S_T 的性质将在下一小节中更详细地讨论. 如果零假设 H_0 成立，则统计量 Z 服从自由度为 $df=T-1$ 的 t 分布. 于是，如果

$$|T|>t(df)_{1-\alpha/2},$$

则在显著水平为 $\alpha\in(0,1)$下拒绝 H_0，其中，$t(df)_{1-\alpha/2}$表示 $t(df)$分布的 $1-\alpha/2$ 分位点.

若对数收益是非正态的，则由中心极限定理知统计量 Z 是渐近正态的. 所以，当$|Z|>z_{1-\alpha/2}$时，零假设被拒绝.

▶**例 1.3.1** 对于图 1-1 所示的 FTSE 对数收益曲线，可以算得 $z=2.340\,558$，超过了 5%显著水平对应的临界值 1.959 964，所以对数收益的均值显著不等于 0，即显著为正. 然而，在 1%的显著水平下该结论是无效的. ◀

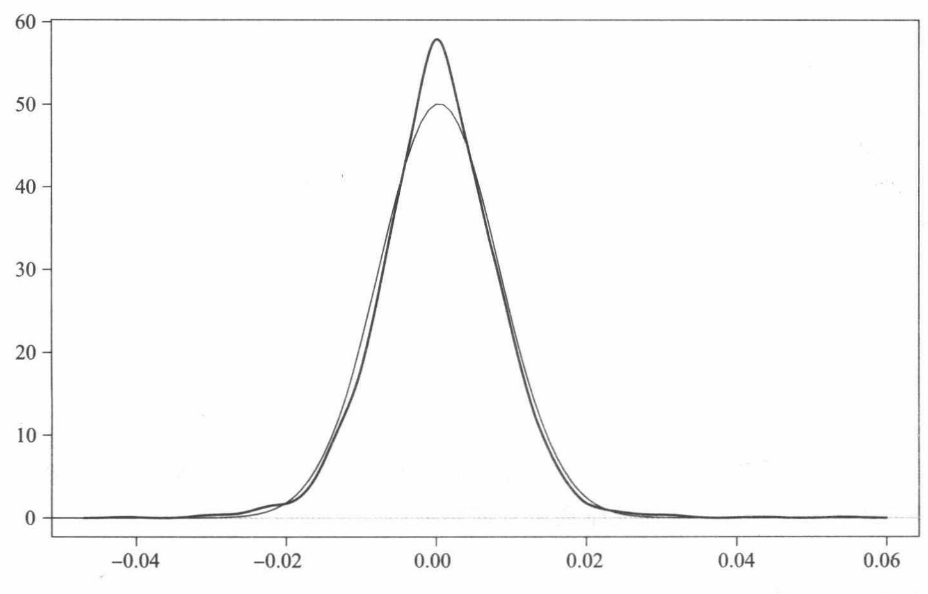


图 1-1 FTSE 日对数收益的基于交叉有效带宽选择的核密度估计曲线

经常，人们也对收益均值的区间估计感兴趣. 假设收益是独立同分布的正态随机变量，若 $L=L(R_1,\cdots,R_T)$，$U=U(R_1,\cdots,R_T)$是样本的函数，且有 $P(L\leqslant\mu\leqslant U)=1-\alpha$ 对任意 $\mu\in\mathbf{R}$ 成立，则称$[L,U]$为收益均值的 $1-\alpha$ **置信区间**. 该置信区间通常有以下表达式

$$L=\overline{R}_T-t(df)_{1-\alpha/2}\frac{S_T}{\sqrt{T}},\quad U=\overline{R}_T+t(df)_{1-\alpha/2}\frac{S_T}{\sqrt{T}},$$

其中自由度 $df=T-1$. 该结论是很容易得到的，事实上，事件 $L\leqslant\mu\leqslant U$ 等价于

$$-t(df)_{1-\alpha/2}\leqslant\sqrt{T}(\overline{R}_T-\mu)/S_T\leqslant t(df)_{1-\alpha/2}.$$

14

因为，

$$\sqrt{T}(\overline{R}_T-\mu)/S_T\sim t(df).$$

所以上面第二个不等式成立的概率为 $1-\alpha$.

在一般情况下，日收益并非正态分布，它受**特征事实**的影响，可能是非对称、巅峰(在 0 附近有更大的密度)和厚尾(相比正态分布)的．这可以从图 1-1 容易地看出．著名的中心极限定理指出，若收益是独立同分布的，且存在四阶矩，则由式(1.8)定义的 Z 统计量是渐近标准正态的.[⊖]于是，一个有效的渐近检验准则为

$$\text{如果}\ |Z|>z_{1-\alpha/2}\ \text{成立，}\quad \text{则拒绝}\ H_0,$$

其中 $z_{1-\alpha/2}=\Phi^{-1}(1-\alpha/2)$表示 $N(0,1)$分布的 $1-\alpha/2$ 分位点．同理，将 L 和 U 等式中的 $t(df)$分布分位点替换为标准正态分布分位点，可以得到 μ 的渐近置信区间.

类似地，可以由中心极限定理构造 μ 的渐近置信区间．此时，事件

15

$$-z_{1-\alpha/2}\leqslant\sqrt{T}(\overline{R}_T-\mu)/S_T\leqslant z_{1-\alpha/2},$$

等价于事件

$$L'=\overline{R}_T-z_{1-\alpha/2}\frac{S_T}{\sqrt{T}}\leqslant\mu\leqslant\overline{R}_T+z_{1-\alpha/2}\frac{S_T}{\sqrt{T}}=U',$$

且当 $T\to\infty$时，其发生的概率趋于 $1-\alpha$．因此，渐近收敛概率为 $1-\alpha$ 的置信区间可以由随机区间$[L',U']$给出.

▶**例 1.3.2** 对于图中的 FTSE 对数收益，可以计算出对数收益均值的 95%渐近置信区间为$[l,u]=[0.000\,070\,2,\ 0.000\,793]$. ◀

大量研究表明，金融收益的统计分析不应作正态性和独立观测的经典假设，因为这些假设通常是不符合事实的．所以，大样本理论构成了金融统计推断的数学核心.

1.3.2 离散程度和风险的度量

均值 $\mu=E(R_t)$给出了分布的位置，它是分布中心的度量，可以确定每个收益 R_t 与其均值的偏离距离$|R_t-\mu|$．称偏离距离平方的均值

$$\sigma^2=\mathrm{Var}(R)=E(R-\mu)^2=E(R^2)-\mu^2$$

为 R 的**方差**，其平方根

$$\sigma=\sigma_R=\sqrt{\mathrm{Var}(R)}$$

称为**标准差**．任意一个存在二阶矩的随机变量 X 均可定义方差和标准差．若 X 和 Y 是互相独立的随机变量，且 $EX^2<\infty$，$EY^2<\infty$，则由

$$\mathrm{Var}(X+Y)=\mathrm{Var}(X)+\mathrm{Var}(Y)$$

可得 $\sigma_{X+Y}=\sqrt{\mathrm{Var}(X+Y)}=\sqrt{\sigma_X^2+\sigma_Y^2}$．在一般情况下，有

$$\mathrm{Var}(X+Y)=\mathrm{Var}(X)+\mathrm{Var}(Y)+2\mathrm{Cov}(X,Y).$$

其中，

$$\mathrm{Cov}(X,Y)=E(X-EX)(Y-EY)$$

称为 X 和 Y 的**协方差**.

在考虑日(对数)收益时，通常称 σ 为(实际)**波动率**．提及收益的波动率时，必须注意

⊖ 在对数收益为一个平稳的鞅差分序列这样一个相当弱的条件下，该结果仍成立.

所使用的时间单位，如年、月、日．**年波动率** σ_{an} 指的是年度收益的标准差，受此启发，对应于时间 τ(以年为单位)的一般波动率为

$$\sigma_{\mathrm{an}} \sqrt{\tau}.$$

16

注意到，若 τ 是正整数且年度对数收益是同分布、互不相关的，则以上公式和 τ 年收益 $R(\tau)$ 的标准差是一致的，这是因为根据对数收益的可加性能够得到 $R(\tau)=\sum_{t=1}^{\tau} R_t'$，其中 R_1'，…，R_τ' 分别表示 τ 年中每年的对数收益，于是

$$\sigma_{R(\tau)}=\sqrt{\sum_{t=1}^{\tau} \operatorname{Var}(R_t')}=\sqrt{\tau}\sigma',$$

这里，σ' 表示年度收益 R_t' 的波动率．年度波动率是由日对数收益的实际波动率决定的，因为一年共有 252 个交易日．年度波动率 σ_{an} 和实际波动率 σ 有如下关系：

$$\sigma_{\mathrm{an}}=\sigma \sqrt{252}.$$

月度波动率则等于 $\sigma_{\mathrm{m}}=\sigma \sqrt{252/12}$.

方差与标准差的估计通常基于前一小节所介绍的“插入准则”．给定收益的一组样本 R_1，…，R_T，则得到**经验方差**或**样本方差**

$$V_T^2=\frac{1}{T} \sum_{t=1}^{T}(X_t-\bar{R}_T)^2.$$

通过计算可得到 $E(V_T^2)=\dfrac{T-1}{T}\sigma^2$，即 V_T^2 不是收益方差的无偏估计．因此，在实际应用中，常常使用如下估计量

$$S_T^2=\frac{1}{T-1} \sum_{t=1}^{T}(R_t-\bar{R}_T)^2.$$

标准差的估计量为上述表达式的平方根，即 $S_T=\sqrt{S_T^2}$.

上面所讨论的各种波动率的估计值可以通过将 σ 替换为 S_T 得到．例如，如果 $R_t s$ 表示的是日对数收益，则年度波动率可以用 $S_T \sqrt{252}$ 来估计．

1.3.2.1 风险值

风险的另一度量——风险值是对金融业风险的本质度量．回忆上节所讲，任何一项投资在[0，h]产生的净收益或亏损(P&L)都是不确定的，因此存在遭受损失的风险．粗略地说，风险值是一种风险度量，它表示在给定一个概率 α 的前提下，可能遭受的最小的损失，这里风险概率 α 的选取取决于自己，一般取为 1%或 5%．设 V_t 为时刻 t 的**盯市价**，即 V_t 是以当前市值为基础的，则收益为 $\Delta V=V_{t+h}-V_t$，若 ΔV 为负表示亏损．现在考虑亏损 $L=-\Delta V$，令 ν 为满足下式的定值：

$$P(L>v)=\alpha.$$

这意味着损失超过 v 的概率是 α．称 v(损失)为在概率水平 α 的**风险值(VaR)**，记为 VaR 或 VaR_α．粗略地说，它是以概率 α 发生损失的大量损失中的最小损失．由定义有

17

$$\mathrm{VaR}_\alpha = F_L^{-1}(1-\alpha),$$

其中，F_L^{-1}表示对应于损失分布的分位数函数. 这意味着：风险值其实是损失分布的$1-\alpha$分位点. 注意，风险值也可以通过 P&L 分布的α分位点来定义，即

$$\mathrm{VaR}_\alpha = -F_{\Delta V}^{-1}(\alpha).$$

一般情况下，风险值都是以日为时间单位计算的. 如果某一天的 1%风险值是 100 000，则该日的值减少 100 000 以上的概率是 1%，遭受损失大于等于 100 000 的概率也是 1%.

由于 VaR 为一个分位数，当然可以通过样本分位数来进行估计. 设L_1，…，L_T表示$[0, h]$内独立同分布的损失变量序列，则有

$$\widehat{\mathrm{VaR}}_\alpha = L_{\lceil n(1-\alpha)\rceil}.$$

统计检验和置信区间的计算可以建立在前面已经介绍的分位数大样本理论基础上，同样，渐近置信区间方法也适用于风险值置信区间的估计.

1.3.2.2 期望损失、低阶偏矩和内在风险测度

VaR 给出了指定概率α下大量损失的最小值，则很自然地会想到损失的平均值，即在给定的时段内考虑收益或损失L的条件期望：

$$S_\alpha(L) = E(L \mid L \leqslant \mathrm{VaR}_\alpha).$$

上式被称为**期望损失**或**条件风险值**. 通过计算可以得到

$$S_\alpha(X) = -\frac{1}{\alpha}\int_0^\alpha F_{\Delta V}^{-1}(x)\mathrm{d}x.$$

因此，$S_\alpha(X)$又被称为**平均风险值**.

显然，在实际中，人们并不关心L中满足条件$l>E(L)$的路径l是否实际发生，而是关心跌价风险，即当日的损失小于$E(L)$. 如果L服从对称分布，那么$P(L<E(L))=P(L>E(L))$，则方差和标准差已经给出了跌价风险的最好度量. 但是，特别是对非对称分布，考虑半方差是有意义的，其定义为

$$E(\min(0, L-EL)^2).$$

证券投资组合的收益通常会和一个基准收益b进行比较. 如果投资组合的收益未胜过基准收益，即若$L\leqslant b$，那么$b-L$就是长期持有该投资组合而遭受的损失，随机变量$(b-L)\mathbf{1}_{\{L\leqslant b\}}$的$m$阶矩

18

$$LP_m(L) = E((b-L)^m \mathbf{1}(L \leqslant b))$$

称为 ***m* 阶下偏矩**(如果它存在的话). 注意到

$$LP_0(L) = P(L \leqslant b)$$

是资产组合收益未超过基准收益的概率，而一阶下偏矩$LP_1(L)$则表示投资组合收益低于基准收益的期望值.

以上所讨论的特征量均对固定时间内投资组合的随机损失赋予了实数值，这些数值也可解释为对风险的量化测度. 问题在于，这些风险测度满足什么性质. 一般地，**风险测度**或**风险函数**ρ是定义在一个充分丰富的随机变量集(这里指随机支付量)$\mathcal{A}$上、取值于实数的函数. 给定一个风险测度ρ，它可以分为**风险支付**(有非负风险)和**可接受支付**(有负

风险).

称风险测度 ρ 是**一致的**，如果它满足下列 4 个条件：

(ⅰ) 对任意的 X，$Y\in\mathcal{A}$，如果 $X\leqslant Y$，则有 $\rho(X)\leqslant\rho(Y)$(单调性)；

(ⅱ) 对任意的 X，$Y\in\mathcal{A}$，$\rho(X+Y)\leqslant\rho(X)+\rho(Y)$(半可加性)；

(ⅲ) 对 $a>0$，$\rho(aX)=a\rho(X)$(正齐次性)；

(ⅳ) 对任意的 $X\in\mathcal{A}$和 $a\in\mathbf{R}$，$\rho(X+a)=\rho(X)-a$(平移不变性)；

有时还会考虑下列条件，

(ⅴ) 如果 $X\overset{d}{=}X'$，那么 $\rho(X)=\rho(X')$(分布不变性).

条件(ⅰ)要求：如果对所有状态 $\omega\in\Omega$，随机支付量均增加，则头寸的风险亦增大. 条件(ⅱ)揭示了风险管理中的一个重要定理：当把两种头寸组合在一起时，二者的风险可能对冲. 标准差测度 $\sigma(X)=\sqrt{\mathrm{Var}(X)}$满足条件(ⅱ)和条件(ⅲ). 首先证明它满足条件(ⅱ)，由不等式

$$\mathrm{Cov}(X,Y)\leqslant\sigma(X)\sigma(Y)$$

得到

$$\begin{aligned}\sigma(X+Y)&=2\sqrt{\mathrm{Var}\left(\frac{X}{2}+\frac{Y}{2}\right)}\\&\leqslant 2\sqrt{\left(\frac{1}{2}\right)^2\sigma_X^2+\left(\frac{1}{2}\right)^2\sigma_Y^2+\sigma_X\sigma_Y}\\&=2\sqrt{\left(\frac{\sigma_X}{2}+\frac{\sigma_Y}{2}\right)^2}\\&=\sigma_X+\sigma_Y.\end{aligned}$$

上式对存在二阶矩且具有任意相关性的随机变量对$(X，Y)$均成立. 由此可知，高斯条件下风险值亦满足条件(ⅱ). 事实上，首先，随机变量 P&L$X\sim N(\mu，\sigma^2)$的风险值可写为

$$\mathrm{VaR}_\alpha(X)=\mu+\Phi^{-1}(\alpha)\sigma.$$

其次，如果二维随机向量$(X，Y)$服从二元正态分布，且其边缘分布为$N(\mu_X，\sigma_X^2)$、$N(\mu_Y，\sigma_Y^2)$，协方差为 γ，则 $X+Y$ 亦服从高斯分布(正态分布)，即

$$X+Y\sim N(\mu_X+\mu_Y,\sigma_{X+Y}^2),\quad \sigma_{X+Y}^2=\sigma_X^2+\sigma_Y^2+2\gamma,$$

且 $X+Y$ 的 α 分位数为

$$\mathrm{VaR}_\alpha(X+Y)=\mu_X+\mu_Y+\Phi^{-1}(\alpha)\sigma_{X+Y}.$$

因此由 $\sigma_{X+Y}\leqslant\sigma_X+\sigma_Y$ 立即得到

$$\mathrm{VaR}_\alpha(X+Y)\leqslant\mathrm{VaR}_\alpha(X)+\mathrm{VaR}_\alpha(Y).$$

但是，对一般的分布，其风险测度并不一定满足条件(ⅱ)，也不一定是一致风险测度，这可能是风险值理论受到批评的主要原因.

条件(ⅲ)是一个比例缩放性质，可以用它来比较以不同货币流表示的风险. 条件(ⅳ)意味着，当给头寸增加了一份支付，为了补偿损失和降低风险，风险测度需要减少

19

同样的量；与之相对的是，从头寸中抽取现金会增大风险．因此，$\rho(X)$可以理解成为了排除风险、将头寸转换为可接受支付而需要的资本．显然，标准差测度不满足条件（ⅳ）．

可以证明，期望损失满足上述所有条件，因此是一致风险测度．更一般地，任何满足条件

$$\rho(X)=\sup_{P\in\mathcal{P}}E_P(-X),$$

的风险测度一定是一致的，其中，$\mathcal{P}$是一个概率测度集，E_P 表示在测度 P 下的期望．对于 $S_\alpha(X)$，集合$\mathcal{P}$为以 $1/\alpha$ 为界的所有概率测度集．

1.3.3　偏度和峰度的度量

偏度是一个常用的概念，它用来度量随机变量与对称的偏离，它为三阶中心矩

$$\mu_3^*=E\left(\frac{R_1-\mu}{\sigma}\right)^3,$$

其中 $\mu=E(R_1)$，$\sigma^2=\mathrm{Var}(R_1)$．如果 $R_1-\mu\xrightarrow{d}\mu-R_1$，则 $\mu_3^*=0$．[⊖]

对给定的一组样本 R_1，…，R_T，μ_3^* 的一个估计量为

20

$$\hat{\mu}_3^*=\frac{1}{T}\sum_{t=1}^{T}\left(\frac{R_t-\overline{R}_T}{S_T}\right)^3.$$

需要指出的是，统计量$\hat{\mu}_3^*$ 对异常离群值有较高的敏感性．度量随机变量与对称的偏离的另一种方法是使用分位数，求出 0.75 分位数与中位数之间的距离以及中位数与 0.25 分位数之间的距离，再求出这两个距离差，最后计算距离差与四分位数距离的比，即

$$\gamma=\frac{[F^{-1}(0.75)-F^{-1}(0.5)]-[F^{-1}(0.5)-F^{-1}(0.25)]}{F^{-1}(0.75)-F^{-1}(0.25)}.$$

γ 对应于样本 R_1，…，R_T 的估计量为

$$\hat{\gamma}_T=\frac{[Q_3-x_{\mathrm{med}}]-[x_{\mathrm{med}}-Q_1]}{Q_3-Q_1}.$$

值得一提的是：对于简单样本，由 Q_1，Q_3, x_{med} 得到的估计量比算术平均估计量更为稳健，即使对来自厚尾分布族的数据，$\hat{\gamma}_T$ 亦给出了偏度的一个可靠度量．

峰度用于测量随机变量的分布与高斯密度的差异，考虑下列四阶中心矩：

$$\mu_4^*=E\left(\frac{R_1-\mu}{\sigma}\right)^4,$$

称 μ_4^* 为**峰度**．由于正态分布的峰度 $\mu_4^*=3$，因此，考虑**超峰度**：

$$\kappa=\mu_4^*-3.$$

在类似于正态分布这种分布中，其超峰度 $\kappa=0$，称为**常峰态**；对 $\kappa\neq0$ 时，若 $\kappa>0$，称为**尖峰态**，它比正态密度函数具有更陡峭的峰值和更薄的尾分布；若 $\kappa<0$，称为**低峰态**，它

⊖　如果 $X\xrightarrow{d}-X$，且函数 f 满足：$f(-x)=-f(x)$，$Ef(X)\in\mathbf{R}$，那么，$Ef(X)=Ef(-X)=-Ef(X)$，从而有 $Ef(X)=0$．

比正态密度函数具有较为平坦的峰值和更厚的尾分布．峰度值和超峰度值 κ 分别可用下列估计量来估计，

$$\hat{\mu}_4^* = \frac{1}{T}\sum_{t=1}^{T}\left(\frac{R_t - \overline{R}_T}{S_t}\right)^4, \quad \hat{\kappa}_T = \hat{\mu}_4^* - 3.$$

1.3.4 分布的估计

前面已经介绍过，在短的时间范围内，收益分布并不具备正态分布的良好性质．幸运的是，收益分布一般都存在密度函数[⊖] f，即分布函数 $F(x)=P(R_1\leqslant x)$可表示为

21

$$F(x) = \int_{-\infty}^{x} f(t)\mathrm{d}t, \quad x \in \mathbf{R},$$

下面讨论如何有效地估计出密度函数 f．注意到 $f(x)=F'(x)$，可以通过差值比来逼近 $f(x)$，即

$$f(x) \approx \frac{F(x+h) - F(x-h)}{2h},$$

其中，$h>0$ 为很小的正数．上式右边项的一个自然的估计方法是使用 T 个历史收益 R_1，…，R_T 的经验分布函数

$$F_T(x) = T^{-1}\sum_{t=1}^{T}\mathbf{1}(R_t \leqslant x)$$

来估计，从而得到 $f(x)$的估计．由

$$\mathbf{1}(R_t \leqslant x+h) - \mathbf{1}(R_t \leqslant x-h) = \mathbf{1}(x-h < R_t \leqslant x+h),$$

可得估计量：

$$x \mapsto \frac{1}{Th}\sum_{t=1}^{T}\frac{1}{2}\mathbf{1}\left(-1 < \frac{R_t - x}{h} \leqslant 1\right), \quad x \in \mathbf{R}.$$

上式右端的每个被加数具有一致形式的密度函数 $K_0(z)=\frac{1}{2}\mathbf{1}(-1<|z|\leqslant 1)$，$z\in\mathbf{R}$，每个点$(x-R_t)/h(t=1,\cdots,T)$均在区间$(-1,1]$上均匀分布．显然，作为 x 的函数，上述密度估计量是非连续的，存在很多“伪跳”．如果把非连续密度 K_0 换成其他函数，则得到 Rosenblatt-Parzen 核密度估计量

$$\hat{f}_{Th}(x) = \frac{1}{Th}\sum_{t=1}^{T}K([R_t - x]/h), \quad x \in \mathbf{R}.$$

称参数 h 为**带宽**，它的选取对估计结果有重要影响．如果选取 h 过小，则模拟曲线中会出现许多诸如局部伪极值点这样的情形，如果带宽值选取过大，虽然曲线会很光滑，但偏差也很大．称 K 为**光滑核函数**，通常取 K 为某个具有有限二阶矩的单峰密度函数，并要求其关于 $x=0$ 对称．表 1-1 给出了一些常用的核函数．

⊖ 对某些金融工具而言，这个假设不成立，因为存在一个交易期，在此期间，价格保持为常数，从而收益为 0.

表 1-1　密度函数的非参估计中一些常用的光滑核函数

核函数	定义	核函数	定义
三角核	$(1-\lvert x\rvert)\mathbf{1}(\lvert x\rvert\leqslant 1)$	Epanechnikov 核	$(3/4)(1-x^2)\mathbf{1}(\lvert x\rvert\leqslant 1)$
余弦核	$(\pi/4)\cos(x\pi/2)$	Biweight 核	$(15/16)(1-x^2)^2\mathbf{1}(\lvert x\rvert\leqslant 1)$
高斯核	$(2\pi)^{-1}\exp(-x^2/2)$	Silverman 核	$(1/2)\exp(-\lvert x\rvert/\sqrt{2})\sin(\lvert x\rvert/\sqrt{3}+\pi/4)$

22

估计量$\hat{f}_{Th}(x)$有如下漂亮的性质：若 K 关于 $x=0$ 对称且有单位方差，那么，对任意的 $m\in\mathbf{R}$，$h>0$，

$$x\mapsto\frac{1}{h}K\left(\frac{x-m}{h}\right),\quad x\in\mathbf{R},$$

是一个均值为 m、标准差为 h 的密度．于是，$\hat{f}_{Th}(x)$是对应于观测值的 T 个密度 $x\mapsto h^{-1}K([x-R_t]/h)(t=1,\cdots,T)$的平均值.

下面进一步讨论密度的核估计量性质，这有助于理解为什么它可以用来估计一般条件下的密度函数．在核估计中，需要讨论如何选取光滑核函数与带宽值．注意到，若 K 是光滑核函数，则$\hat{f}_{Th}(x)$是一个密度函数；其次，$\hat{f}_{Th}(x)$保留了 K 的许多好的光滑性质．特别地，可以利用$\hat{f}'_{Th}(x)$去估计 $f'(x)$．若收益 $R_1,\cdots,R_T$ 是独立同分布的样本，则有

$$E(\hat{f}_{Th}(x))=\int\frac{1}{h}K\left(\frac{z-x}{h}\right)f(z)\mathrm{d}z=(K_h\star f)(x),$$

其中，$K_h(z)=h^{-1}K(z/h)$是修正后的核密度，$\star$表示卷积算子．由此可知，Rosenblatt-Parzen 估计量不是 f 的无偏估计量，其偏差为

$$b_h(x)=E(\hat{f}_{Th}(x))-f(x)=(K_h\star f)(x)-f(x).$$

但是，由 Bochner 引理(引理 A. 2. 1)知，当 $h\to 0$ 时，卷积$(K_h\star f)(x)$收敛于 $f(x)$．因此，带宽应该选择为样本容量 T 的递减函数．在独立同分布(i. i. d.)的假设下，易证$\hat{f}_{Th}(x)$的方差为

$$\sigma_{Th}^2(x)=\mathrm{Var}(\hat{f}_{Th}(x))=\frac{1}{Th}[(K_h^2\star f)(x)-(K_h\star f)^2(x)],$$

其中 $K_h^2(z)=h^{-1}K^2(z/h)$，$z\in\mathbf{R}$．同样由 Bochner 引理知，上式方括内的表达式收敛于一个有限常数，因此，当 $Th\to\infty$时，$\hat{f}_{Th}(x)$的方差与 $1/Th$ 同阶，所以方差收敛于 0．现考虑**均方误差**(MSE)：

$$\mathrm{MSE}(\hat{f}_{Th}(x);f(x))=E(\hat{f}_{Th}(x)-f(x))^2,$$

它可以分解为方差 $\sigma_{Th}^2(x)$和平方偏差 $b_h^2(x)$两项之和，即

$$\mathrm{MSE}(\hat{f}_{Th}(x);f(x))=\sigma_{Th}^2(x)+b_h^2(x).$$

可以发现，对任意满足

23

$$h\to 0\quad 且\quad Th\to\infty$$

的带宽值，均方误差 MSE 都收敛于 0.

为了作进一步的讨论，引入下列记号：称核函数 K 为 ***r* 阶核**，如果它满足

$$\int z^j K(z)\mathrm{d}z = \begin{cases} 1, & j=0, \\ 0, & j=1,\cdots,r-1, \\ c\neq 0, & j=r. \end{cases}$$

例如，$K(x)=\left(\frac{9}{8}-\frac{15}{8}x^2\right)\mathbf{1}(|x|\leqslant 1)$，$x\in\mathbf{R}$ 是 4 阶核. 如果密度函数 f 是 r 阶可微的，则易得下列展式：

$$E(\hat{f}_{Th}(x)) = f(x)+h^r f^{(r)}(x)\frac{(-1)^r}{r!}\int u^r K(u)\mathrm{d}u + o(h^r),$$

$$\mathrm{Var}(\hat{f}_{Th}(x)) = \frac{1}{Th}f(x)\int K^2(z)\mathrm{d}z + o(1/Th),$$

从而可得 MSE 的下列展式：

$$\mathrm{MSE}(\hat{f}_{Th}(x);f(x)) = \frac{f(x)R(K)}{Th}+h^{2r}[f^{(r)}(x)]^2 M_r^2 + o(h^{2r}+1/Th),$$

其中

$$M_r = \frac{(-1)^r}{r!}\int u^r K(u)\mathrm{d}u$$

而

$$R(g) = \int g^2(x)\mathrm{d}x$$

表示 L_2 空间中函数 g 的**粗糙度**. 以上展式表明，高阶核可以导出均方误差的阶，此例为 $O(h^{2r})$.

如果选取带宽 h 使得上面展式的主部值，即函数

$$h\mapsto \frac{f(x)R(K)}{Th}+h^{2r}[f^{(r)}(x)]^2 M_r^2,\quad h>0$$

达到最小，称这样的 h 为**局部渐近最优带宽**. 易证，最优带宽可表示为

$$h^*(x) = h_T^*(x) = \left(\frac{f(x)R(K)}{2rM_r^2[f^{(r)}(x)]^2 T}\right)^{1/(2r+1)}.$$

特别地，一个二阶核的最优带宽的阶为 $O(T^{-1/5})$，这个结论有助于局部带宽的选取. 在具体应用该方法时，需要不断地用 $f(x)$和导数 $f^{(r)}(x)$的试估计量进行估计.

然而，实践中用得更广泛的方法还是建立在**积分均方误差**(IMSE)基础上的方法，令

$$\mathrm{IMSE}(\hat{f}_{Th};f) = \int \mathrm{MSE}(\hat{f}_{Th}(x);f(x))\mathrm{d}x = \int E(\hat{f}_{Th}(x)-f(x))^2\mathrm{d}x.$$

24

当 $r=2$ 时，有下列展式：

$$\mathrm{IMSE}(\hat{f}_{Th};f) = \frac{R(K)}{Th}+\frac{1}{4}h^4 M_2^2\int[f^{(2)}(x)]^2\mathrm{d}x + o(h^4+1/Th).$$

略去余项得到**渐近积分均方误差**(AIMSE)为

$$\mathrm{AMISE}(h) = \frac{R(K)}{Th}+\frac{1}{4}h^4 M_2^2\int[f^{(2)}(x)]^2\mathrm{d}x,$$

将其视为带宽 h 的函数，不难看出，最小化 AIMSE 的最优带宽 h_{opt} 为

$$h_{\mathrm{opt}} = C_0 T^{-1/5},$$

其中

$$C_0 = M_2^{-2/5} R(K)^{1/5} \left[\int f^{(2)}(x)]^2 \mathrm{d}x \right]^{-1/5}.$$

遗憾的是，常数 C_0 是未知的. 但当以均值为 0，方差为 σ^2 的标准正态分布作为参考模型的时候，使用正态经验法则可以求出该常数. 若采用正态核，则有如下最优带宽：

$$h_{\mathrm{opt}}^{*} = (4\pi)^{-1/10}[(3/8)\pi^{-1/2}]^{-1/5}\sigma \cdot T^{-1/5} \approx 1.06\sigma T^{-1/5}.$$

实践中，参数 σ 一般用样本标准差代替.

显然，上述方法有一个明显的缺陷，在估计渐近最优带宽时，它是针对一个固定的参考分布，但是核估计的目的是针对任意的未知分布. 所以需要引入克服了上述局限的方法. 下面简要地介绍一下目前广泛使用的无偏和有偏最小二乘交叉检验法.

最小二乘无偏交叉检验法是使平方误差的积分这个非参数估计量达到最小而得到的带宽，因此针对所有 x 而不是针对固定的某个 x 提供最优带宽. 因为

$$\int [\hat{f}_{Th} - f(x)]^2 \mathrm{d}x = \int \hat{f}_{Th}^2(x)\mathrm{d}x - 2\int \hat{f}_{Th}(x) f(x)\mathrm{d}x + \int f(x)^2 \mathrm{d}x,$$

则最小化积分均方误差等价于最小化上式右边的前两项. 如果 $R \sim f$ 且与 R_1, …, R_T 独立，E_R 表示关于 R 的期望，则有

25

$$\int \hat{f}_{Th}(x) f(x)\mathrm{d}x = E_R(\hat{f}_{Th}(R)),$$

所以，可以利用下式

$$\hat{f}_{T,-i} = \frac{1}{(T-1)h} \sum_{t=1, t\neq i}^{T} K\left(\frac{R_t - R_i}{h}\right).$$

去估计 $E_R(\hat{f}_{Th}(R))$. 称上式为 $f(x_i)$ 的**留一法估计**. 至于第一项，通过代入核密度估计得到

$$\int \hat{f}_{Th}^2(x)\mathrm{d}x = \frac{1}{T^2 h^2} \sum_{t=1}^{T} \sum_{s=1}^{T} \int K\left(\frac{R_t - x}{h}\right) K\left(\frac{R_s - x}{h}\right)\mathrm{d}x$$

$$= \frac{1}{T^2 h} \sum_{t=1}^{T} \sum_{s=1}^{T} (K\star K)\left(\frac{R_t - R_s}{h}\right).$$

利用这两个估计量得到：最小二乘交叉检验法是最小化如下目标函数：

$$\mathrm{UCV}(h) = \frac{1}{T^2 h} \sum_{t=1}^{T} \sum_{s=1}^{T} (K\star K)\left(\frac{R_t - R_s}{h}\right) - \frac{2}{T(T-1)h} \sum_{s=1}^{T} \sum_{t=1, t\neq s}^{T} K\left(\frac{R_t - R_s}{h}\right),$$

这只能通过数值解法求得. 从形式上看，UCV(h)表达式中两项的期望与 IMSE 前两项吻合. 可以证明，最小化 CV(h)渐近等价于最小化

$$B_1 h^4 + \frac{R(K)}{Th},$$

其中

$$B_1 = \frac{M_2^2}{4}\left\{ \int [f^{(2)}(x)]^2 \mathrm{d}x \right\}.$$

由此容易看出，后者的最小化解和 IMSE 的最小化解是一致的. 而且，可以证明，当 $T \to \infty$ 时，依概率有

$$\frac{h_{\mathrm{LCV}} - h_{\mathrm{opt}}}{h_{\mathrm{opt}}} \to 0,$$

它是这种方法的一个更加强大的工具.

有偏最小二乘交叉检验是使渐近均方误差(AMISE)的另一个估计最小化. 前面提到

$$\mathrm{AMISE}(h) = \frac{R(K)}{Th} + \frac{1}{4} K_2^2 h^4 R(f'').$$

解得的最优化带宽值为

$$h_0 = \left(\frac{R(K)}{M_2^2 TR(f'')} \right)^{1/5}.$$

26

对上式中的唯一未知量 $R(f'')$，自然想到用 $R(\hat{f}''_T)$来估计，其中 $\hat{f}''_T$ 表示核估计 $\hat{f}_T$ 的二阶导数，它满足

$$E(R(\hat{f}''_T)) = R(f'') + \frac{R(K'')}{Th^5} + O(h^2).$$

还可以估计正偏差得到更好的结果，它得到估计量 $R(\hat{f}''_T) - \dfrac{R(K'')}{Th^5}$. 注意到

$$R(\hat{f}''_T) = \frac{R(K'')}{Th^5} + \frac{2}{T^2 h^5} \sum_{1 \leqslant s < t \leqslant T} \varphi\left(\frac{X_t - X_s}{h} \right),$$

其中

$$\varphi(x) = \int K''(u) K''(u+x) \mathrm{d}u, \quad x \in \mathbf{R},$$

从而得到有偏最小二乘交叉检验函数为

$$\mathrm{BCV}(h) = \frac{R(K)}{Th} + \frac{K_2^2}{2T^2 h} \sum_{1 \leqslant s < t \leqslant T} \varphi\left(\frac{X_t - X_s}{h} \right),$$

这就是所需的最小化解.

图 1-1 给出了 FTSE 从 1991 年到 1998 年每日对数收益的核密度估计量，其中的带宽是使用有偏最小二乘交叉检验法得到的.

1.3.5 正态性检验

资产收益的分布一般是非正态的，特别是来自短时间段上的样本，例如每日收益或当天收益. 为了检验收益是否服从正态分布，有许多的统计检验方法可用. 在本小节中，将介绍几种实践中使用最广泛的检验方法.

设 $R_1, \cdots, R_T$ 为收益的一个 i. i. d. 样本，分布函数为 F. 我们的目标是检验原假设 F 为正态分布，即

$$H_0 : F \in \{\Phi_{\mu,\sigma^2} : \mu \in \mathbf{R}, \sigma^2 > 0\},$$

备择假设为

$$H_1 : F \notin \{\Phi_{(\mu,\sigma^2)} : \mu \in \mathbf{R}, \sigma^2 > 0\}.$$

注意，H_1 成立意味着对任意的 $\mu\in\mathbf{R}$ 和 $\sigma^2>0$，至少存在一个 $x\in\mathbf{R}$，使得 $F(x)\neq\Phi_{(\mu,\sigma^2)}(x)$.

正态性检验的 Jarque-Bera 检验统计量为

$$J_T=T\Big(\frac{\hat{\mu}_3^2}{6}+\frac{(\hat{\mu}_4-3)^2}{24}\Big),$$

其中 $\hat{\mu}_3$ 是样本偏度，$\hat{\mu}_4$ 是样本峰度. 当 $T\to\infty$ 时，J_T 渐近服从 $\chi^2(2)$ 分布，所以，如果 $J_T>\chi^2(2)_{1-\alpha}$，则拒绝 H_0. 该检验只适用于大样本的检验. 注意，在原假设条件下，Jarque-Bera 检验统计量度量了样本偏度和样本峰度与它们理论值的偏离程度.

27

另一类检验是基于以下想法：如果原假设成立，则用样本估计量 $\hat{\mu}_T$ 和 S_T^2 来估计参数 μ 和 σ^2，相应地，分布函数的估计量为 $\Phi_{(\hat{\mu}_T,S_T^2)}(x)$. 如果备择假设成立，则用经验分布函数，即

$$\hat{F}_T(x)=\frac{1}{T}\sum_{t=1}^{T}\mathbf{1}(R_t\leqslant x),\quad x\in\mathbf{R},$$

作为分布函数 $F(x)$ 的估计量. 在对 $F(x)$ 没有任何形状假设的前提下，它是 $F(x)$ 的一致估计量. 现通过计算二者的最大偏差来比较这两个估计量，为此，引入下列 **Lilliefors 检验统计量**

$$L=\sup_{t\in\mathbf{R}}\left|\hat{F}_T(t)-\Phi_{(\hat{\mu}_T,S_T^2)}(t)\right|.$$

遗憾的是，L 的渐近分布并不是目前已知的某个标准分布. 但对 5% 的显著性水平，可以通过比较 L 值与临界值 $0.805/\sqrt{T}$ 的大小实现，其计算过程一般是在标准统计软件中完成的.

有时候，给定了 $\mu_0\in\mathbf{R}$ 和 $\sigma_0^2>0$，要作

$$\text{原假设 } H_0:F=\Phi_{(\mu_0,\sigma_0^2)}(x)\longleftrightarrow\text{备择假设 } H_1:F\neq\Phi_{(\mu_0,\sigma_0^2)}(x)$$

的假设检验. 在这种情况下，可取如下统计量：

$$KS=\sup_{t\in\mathbf{R}}\left|\hat{F}_T(t)-\Phi_{(\mu_0,\sigma_0^2)}(t)\right|.$$

这种检验方法称为 **K-S(Kolmogorov-Smirnov)检验**.

1.4 金融工具

在介绍本节的主要内容之前，需要介绍一些术语和基本的金融衍生产品. 从经济学的角度看，交易是买卖双方达成的协议，约定在某个交易日买卖一定数量的某种资产. 买方持有资产的**多头头寸**，卖方持有**空头头寸**. 每一项交易必然涉及支付，对交易的任意一方，作如下的符号规定：交易者收到支付(即收到付款或报酬)，记为正；反之记为负.

1.4.1 未定权益

很多金融工具的支付依赖于其他工具或者资产，如股票、股票指数、石油、能源价格或者商品等等，这些被依赖的工具或资产统称为该工具的**标的资产**(原生资产). 现代金融市场上甚至可以买卖以天气质量为标的物的金融工具.

衍生品和期货是为了对实体经济中商品和服务的生产分配进行风险规避而产生的，事实上，市场上也的确存在这类需求. 然而，现在越来越多的衍生品交易仅仅出于投机目的. 在一定程度上，市场需要投机者作为风险对冲的对手存在，但是有些市场被过多的投

28

机者操纵而产生了大量的经济泡沫．例如，在农产品市场上，自从2005年以来，受极端天气和市场需求攀升的影响，发展中国家的食品价格出现了前所未有的大幅度上涨，导致了饥荒和社会的不稳定，而一些不道德的投机行为在其中扮演了重要的不光彩角色．

称支付依赖于其他资产的金融工具为**未定权益**，它的严格数学定义将在后面给出．如果未定权益的标的物是类似于股票的证券，则称之为**衍生资产**．下面介绍一些最重要的衍生产品和相关的金融工具及合约．

1.4.2 现货合约与远期合约

定义 1.4.1 **现货合约**是指双方签订的在当天以现货价格(记为 S_t)买卖一项资产的协议．在以后的内容中，均假设 $t=0$ 表示交易初始时刻，T 表示交易结束时刻．**远期合约**是指双方约定在未来某确定时间以确定价格买卖一项资产的协议，其交易价格一般以合约规定为准．远期合约的多头持有者可以在交割日期 T 以执行价格 K 买入资产，而 K 与远期价格 F 保持一致．远期合约的多头持有者的到期收益为 S_T-K，而空头收益为$K-S_T$．

现货合约的交易市场称为现货市场．远期合约在市场外交易，一般发生在金融机构及客户(如企业和私人投资者)之间．在 $t=0$ 时刻双方没有现金支付．在交割日 T，合约卖方需要把资产转移给买方，远期交易才算完成，但在实际中，双方往往以现金交割．若到期日现价 S_T 高于执行价格 K，合约多头赚取 S_T-K 的支付，该项额外支付使得多头愿意以远期价格 $F=K$ 买入资产，因为 $-S_T+(S_T-K)=-K$．但是当到期日的实际价格低于 K 时，他需要支付价差给卖方，双方同样存在着额外支付，所以资产买卖的净价是交割价格．

1.4.3 期货合约

定义 1.4.2 **期货**通常是在交易所交易的标准化远期合约．

例如，美国纽约商业交易所轻质原油期货是标准化的每月1000桶的实物交割合约．交易所的标准化管理允许市场参与者活跃地互换合约．因此，与高度特殊化、无法流通的远期合约相比，期货是流动性很强的金融工具．交易所规范了资产交易的每个细节，如一项合约将有多少单位的资产交易(即合约股数)，交割日期以及资产交割地点、交割方式等．对许多期货合约来说，实物交割往往不太可能或不太方便，因此代以现金结算．合约双方的等价现金支付是初始化的．在到期日之前，期货合约可以在任何时刻流通买卖．与期货合约相关的价格是期货价格，在每个交易日都会确定一个结算价格，常常是该交易日的收盘价，结算价格用来确定投资者必须提供的保证金．投资者需要在保证金账户中存放足额的保证金，当签订合约时，需要支付初始保证金．在每个交易日，该账户采用逐日盯市制，不断调整账户损失和收入．当期货价格上升，合约多头盈利额恰好等于空头损失额，空头投资者的经纪人减少其保证金账户余额，交易所将该项资金转移给对方经纪人，对方账户余额增加．该制度称为逐日清算．若保证金账户余额低于最低保证金，投资者会收到追加保证金的要求，并且需履行该义务，否则经纪人将其头寸清仓，使得合约失效．

1.4.4 期权

期权是给予多头持有者(即期权买方)一项选择权而非义务的协议，多头投资者可以在未来满足某个条件下以确定的价格买卖标的资产. 目前，交易市场上出现了众多不同种类的期权，下面给出一些基础的期权定义.

定义 1.4.3(欧式看涨/看跌期权，基价，到期日) **欧式看涨期权**赋予持有者在确定日期 T(称为**到期日**)以确定价格 K(称为**敲定价格**或**基础价格**)买入某种标的资产的权利；**欧式看跌期权**的所有者拥有在到期日以敲定价格出售标的资产的权利. 用 S_t 表示标的资产在时刻 $t\in[0, T]$时的价格，欧式看涨期权在 t 时刻的价格(即期权的公平价值)记为 $C_e(S_t, K, t, T)$，相应地，记 t 时刻看跌期权的价格为 $P_e(S_t, K, t, T)$. $T-t$ 则称为**到期期限**.

期权常常采用现金结算. 这意味着买方并不实际买入标的资产，而是向对方收取其盈利的现金，该现金为标的资产的敲定价和到期日市场价的差. 记 $C(S_t, K, t, T)$为一项期权在时刻 t 的价格，则在时刻 T，它的值由下式表达：

$$s\mapsto C(s,K,T,T),\quad s\in[0,\infty).$$

若 $S_T>K$，欧式看涨期权所有者行权，盈利 S_T-K. 因此：

$$C_e=C_e(S_T,K,T,T)=\begin{cases}S_T-K, & S_T>K,\\0, & S_T\leqslant K,\end{cases}$$

上式等价于：

$$C_e=\max(0,S_T-K)=(S_T-K)^+.$$

类似地，对于欧式看跌期权有：

$$P_e=P_e(S_T,K,T,T)=\max(0,K-S_T)=(K-S_T)^+.$$

期权的**内在价值**是指期权行权时的正现金流入. 对看涨期权，它等于$(S_T-K)\mathbf{1}(S_T>K)$，看跌期权则为$(K-S_T)\mathbf{1}(S_T<K)$. 若期权的内在价值为正(即看涨期权对应 $S_t>K$，看跌期权对应 $S_t<K$)，称期权处于**实值状态**(或**价内期权**)，若内在价值为 0，称之为**价外期权**. 价格比 K/S 称为**货币性**.

30

▶**例 1.4.4**(资产组合保值) 欧式看跌期权可以用来解决例 1.1.1 中的问题. 假设养老基金打算购买一个股票投资组合(通常称为股票篮子)，并设股票的当前价为 $S_t=110$. 进一步假定该基金可以购买该股票篮子的欧式看跌期权. 设基金最多可以承受 10 单位货币的股票跌价风险，看跌期权的敲定价格设为 100.

养老基金的资产组合中包括一只股票和一份欧式看跌期权，考虑在到期日 T 的价值. 若 $S_T>100$，期权处于虚值状态，也就是说，其价值为 0，则该组合的价值为 S_T；如 $S_T\leqslant 100$，期权收益为 $100-S_T$，资产组合的价值为 $V_T=S_T+(100-S_T)=100$. 因此，无论股价如何变化，养老基金的损失都不会超过 10 单位货币. ◀

1.4.5 障碍期权

看涨障碍期权的价值依赖于标的资产的价格能否达到某个阈值，该阈值称为障碍.

障碍期权又可分为敲出和敲入两种. 若障碍被"击穿"，则敲出期权失效，而敲入期权生效.

定义 1.4.5 设欧式障碍期权到期日为 T，设障碍为 B，$B<S_0$ 且 $B<K$，K 为敲定价格. 如果仅当

$$S_t > B \text{ 对所有 } 0 \leqslant t \leqslant T \text{ 均成立},$$

期权所有者才拥有在 T 时刻执行期权的权利，则称该欧式期权为向下敲出欧式期权. 如果仅当

$$S_t < B \text{ 对所有 } 0 \leqslant t \leqslant T \text{ 均成立},$$

期权所有者才拥有在 T 时刻执行期权的权利，则称该欧式期权为向上敲出欧式期权. 对应地，仅当

$$S_t \leqslant B \text{ 对某个 } t \in [0,T] \text{ 成立},$$

期权所有者才拥有在 T 时刻执行期权的权利，则称该欧式期权为向下敲入欧式期权. 仅当

$$S_t \geqslant B \text{ 对某个 } t \in [0,T] \text{ 成立},$$

期权所有者才拥有在 T 时刻执行期权的权利，则称该欧式期权为向上敲入欧式期权.

与欧式期权不同的是，只要满足行权条件，美式期权允许持有人在到期前任意时刻买入标的资产.

障碍期权是**路径依赖期权**的典型例子. 所谓路径依赖期权，是指其支付和价值均依赖于有效期内标的资产的价格路径 S_t，$0 \leqslant t \leqslant T$ 的期权.

定义 1.4.6 **美式平均价格看涨期权**是执行时收益由下列函数确定的美式期权：

$$\max(0, \overline{S_t} - K), \quad t = 1, \cdots, T,$$

其中 K 为执行价，而

$$\overline{S}_t = \frac{1}{t}\sum_{i=1}^{t} S_i, \quad t = 1, \cdots, T,$$

31

表示标的资产的平均价格. 美式**敲定看涨期权**是收益函数为 $\max(0, S_t - \overline{S_t})(t=1, \cdots, T)$ 的美式看涨期权，它们的执行价格在期权执行时才确定. 对应的欧式期权的收益函数分别为 $\max(0, \overline{S_T} - K)$ 和 $\max(0, S_T - \overline{S_T})$.

1.4.6 金融工程

通过组合不同的金融工具，特别是衍生产品，人们可以进行有趣的盈利分析. 例如，**多头跨式**指同时持有欧式看涨期权和欧式看跌期权的多头头寸，两种期权具有相同的标的资产和相同的到期时间. 若股价上涨，看涨期权获利；若股价下跌，看跌期权获利. 通过这种方式，投资者可以构造出始终盈利的头寸模式，与股价涨或跌无关.

可以证明，衍生产品和未定权益的公平价格 π 可以通过某个概率测度下的贴现收益 C^* 的期望值 $E^*(C^*)$ 来计算. 该方法同样适合证券组合进行定价. 设该组合包含 n 个不同头寸，对应的贴现收益分别为 C_1^*，$\cdots$，C_n^*，设持有的头寸数量分别为 x_1，$\cdots x_n$，因为期望具有线性性质，所以可以得到该证券组合的公平价格为

$$E^*\left(\sum_{i=1}^{n} x_i C_i^*\right) = \sum_{i=1}^{n} x_i \pi_i,$$

其中，$\pi_i = E^*(C_i^*)$表示第 i 个权益的公平价格.

在金融工程中，经常需要构造衍生产品组合进行盈利，例如，为了同时对冲风险和创造获利机会，或者作为一项复合金融商品出售给投资者．如果复合产品 Z 由资产组合构成，即 $Z = \sum_{i=1}^{n} x_i C_i$)，则由上面公式可以得到 Z 的公平价格．由于该复合产品的风险来源于标的工具 C_1，…，C_n 的风险，而影响这些风险的因素包括利息率、标的资产价格变化、标的资产种类变动或者是发行方的失误等．对顾客来说，潜在的投资组合是未知的，这容易造成投资者对该商品的风险作出错误的评估.

1.5 期权定价基础

本节主要介绍期权定价的基本思想和原理．期权定价可以在仅有一份资产和一份欧式看涨期权的金融市场框架下进行简化分析，这已是不争的事实．所以本节首先介绍如何确定仅有一份未定权益的期权的公平价格，然后将其推广到投资组合中去.

1.5.1 无套利原理

无套利原理是指在一个理想的金融市场中，公平的价格不允许出现无风险套利，也就是说，“天下没有免费的午餐”．例如，如果纽约的商品价格远远高于伦敦的价格或者某项债券的价格低于其未来支付的公平价值，则存在套利机会．下面给出套利机会的严格数学定义.

32

定义 1.5.1 套利机会是指在时刻 $t=0$ 初始支付为 x_0，在 $t=1$ 时刻的收益为随机变量 X_1，并且满足

$$x_0 \leqslant 0 \quad (\text{无成本})$$

和

$$X_1 \geqslant 0 \quad P\text{-a.s.} \quad \text{及} \quad P(X_1 > 0) \geqslant 0.$$

▶**例 1.5.2** 利用无套利原理给出远期合约的公平价值，即在 $t=0$ 时刻对其作无套利定价．考虑连续复利的情形，其无套利远期价格为 $F_0 = S_0 e^{rT}$．如果 $F_0 > S_0 e^{rT}$，在这种情况下，期权出售方(远期合约空头)通过在 $t=0$ 借入 S_0 用于买入标的资产，则可赚取无风险利润，因为在到期日 T，它以交割价 K 卖出资产，偿还 $S_0 e^{rT}$ 后还赢利 $F_0 - S_0 e^{rT} > 0$．若 $F_0 < S_0 e^{rT}$，则买方(远期合约多头)可以预先卖空标的资产并把所得货币存入银行，至到期日 T，银行资金增值为 $S_0 e^{rT}$，而他只需支付 F_0 便可对冲资产空头，尚余 $S_0 e^{rT} - F_0$．远期合约的一个重要且有趣的结论是：远期价格不依赖于标的资产到期日的价格. ◀

无套利定价原理也适用于欧式看涨期权与欧式看跌期权的定价，其主要思想是构造一个与欧式看涨期权有相同到期日 T 以及敲定价格 K 的投资组合．设投资者买入一份股票，卖出一份面值为 K 的零息债券，则 T 时刻的价值为 $S_T - K$．又假设在该组合中还购买了

一份看跌期权，在 T 时刻，若 $S_T>K$，看跌期权价值为 0；若 $S_T\leqslant K$，则看跌期权价值为 $K-S_T$. 于是，当 $S_T>K$ 时，T 时刻资产组合的总价值变为 S_T-K；当 $S_T\leqslant K$ 时，T 时刻资产组合的总价值为 0. 其在 $t=0$ 时刻的价值为：

$$\pi(P_e)-K\mathrm{e}^{-rT}+S_0$$

显然该投资组合在 $t=0$ 时刻的价值为一份欧式看涨期权的公平价值，所以有下列平价关系

$$\pi(C_e)=\pi(P_e)-K\mathrm{e}^{-rT}+S_0.$$

存在套利机会与无套利原理是对立的. 存在套利机会意味着金融工具的当前价格和它未来的支付流不一致. 许多经济学家认为，实际金融市场中最多是暂时性存在套利机会的，套利是由市场主体发现的，迅速交易获取套利会导致该机会很快消失. 例如，如果某金融工具的实际定价过低，投机者会成为资产的多头并且等待其价格上升，直到无风险利润消失. 下面将看到，无套利原理是公平定价中一个简单而又强大的工具.

1.5.2 风险中性定价

对未来的随机支付 X 的估值与对风险的偏好有关，它可以由测度空间里的一个概率测度来表示. 其核心问题是：是否它是一个确定性支付，即存在某个常数 x_0，使得对任意 $\omega\in\Omega$，都有 $X(\omega)=x_0$？或者风险支付给了 $\{X>x_0\}$ 发生的机会，或者有发生事件 $\{X<x_0\}$ 的风险.

33

为了简化讨论，假设未来支付的不确定性是由支付波动率(即方差的平方根)引起的. 考虑两个具有相同收益均值的投资，**风险厌恶者**选择其中具有较小方差的策略，相比之下，**风险中性**的投资者没有特别偏好，他将忽略方差带来的影响.

在由风险中性投资者组成的风险中性世界里，所有人都同样看待投资收益. 记 P^* 为风险中性概率测度，在风险中性测度下，一份股票被认为是有意义的无风险投资，当且仅当它有比银行更高的无风险期望收益. 用 S_t 表示股票在 t 时刻的股价，其随机收益为 R，设 $t=0$ 时股价 S_0 是已知的常数，则 $t=1$ 时刻的随机价格为

$$S_1=S_0(1+R).$$

在风险中性世界里，支付价格为

$$E^*(S_1)=S_0(1+E^*(R)).$$

若无特殊说明，符号 E^* 均表示在风险中性测度 P^* 下求期望. 如选择把初始资本 S_0 存入银行账户，则 $t=1$ 时的资金为 $S_0(1+r)$. 根据无套利原理，$E^*(S_1)$和 $S_0(1+r)$必须一致，所以有

$$E^*(S_1)=S_0(1+r)\quad\Leftrightarrow\quad E^*\left(\frac{S_1}{1+r}\right)=S_0.$$

则股票的风险中性价格可以通过求概率测度 P^* 下的期望得到. 不禁要问，可以通过以上方程求出 P^* 吗？

为了快速入门，首先考虑一个单期模型. 假设金融市场上仅有一只股票以及以该股票为标的资产的欧式看涨期权，为了使模型尽可能地简化，设股价服从上涨或下跌两种情况的二叉树假设. 在这种情况下，可以选择样本空间 $\Omega=\{+,-\}$ 作为金融市场未来可能发

生的情况，在 Ω 中配备 σ 代数为幂集. 真实概率测度 P 由 $P(\{+\})=p$，$p\in(0,1)$唯一确定. 注意到，在 P 的定义中，已经排除了 $p=0$，$p=1$ 的特殊情况. 则股价模型为：

$$S_1(\omega)=\begin{cases}S_0u, & \omega=+,\\ S_0d, & \omega=-,\end{cases}$$

其中，常数 u 和 d 分别表示上升因子和下降因子，且 $0<d<1+r<u$ 成立. 则欧式看涨期权的未定权益为

$$C_e=\begin{cases}S_1-K, & S_1>K,\\ 0, & S_1\leqslant K.\end{cases}$$

34

为了避免繁琐的讨论，假设敲定价格 K 满足 $S_0d<K<S_1u$.

在上述简单模型中，风险中性概率测度 P^* 由 $p^*=P^*(\{+\})$唯一确定. 现在，风险中性定价公式 $E^*(S_1)=S_0(1+r)$等价于

$$p^*S_0u+(1-p^*)S_0d=S_0(1+r),$$

由上述方程唯一地解得

$$p^*=\frac{1+r-d}{u-d}.$$

这意味着，如果给定了方程参数 r，d 和 u，则可以确定 P^*. 根据风险中性定价原理，随机支付 X_1 在时刻 $t=1$ 的公平价格可由下式计算，即

$$\pi(X_1)=E^*(X_1/(1+r)).$$

特别地，对股票欧式看涨期权，有

$$\pi(C_e)=p^*\frac{S_0u-K}{1+r}.$$

▶**例 1.5.3** 回顾例 1.1.3 和例 1.1.4，设其中原油价格或者上涨 10%，或者下跌 10%. 即 $u=1.1$，$d=0.9$，又设无风险利率 $r=0.01$. 则风险中性概率测度 P^* 对应的 p^* 值为

$$p^*=\frac{1+r-d}{u-d}=\frac{1.01-0.9}{0.2}=0.55,$$

由此得到期权的风险中性价格为

$$E^*(C_e/(1+r))=\frac{10}{1.01}\times 0.55=5.445\,545,$$

恰好等于原油交易商计算得到的低价极限. ◀

现将模型推广到股价的三叉树模型中去. 令 $\Omega=\{+,\circ,-\}$，假设给定三个常数因子 $d<m<u$，且在 $t=1$ 时，股价满足

$$S_1(\omega)=\begin{cases}S_0u, & \omega=+,\\ S_0m, & \omega=\circ,\\ S_0d, & \omega=-.\end{cases}$$

现在，风险中性概率测度 P^* 由 p_1^*，p_2^*，$p_3^*\in[0,1]$决定，其中 $p_1^*+p_2^*+p_3^*=1$. 在这个模型中，由定价公式 $E^*(S_1)=S_0(1+r)$得到

$p_1^* u + p_2^* m + (1 - p_1^* - p_2^*)d = (1+r) \quad \Leftrightarrow \quad p_1^*(u-d) + p_2^*(m-d) = (1+r) - d.$
但该方程有无穷多个解，而对应于 $p_2^* = 0$ 的特解就是前面的二叉树模型. 一般情况下，方程的通解可以用参数 p_2^* 表示，有

$$p_1^* = \frac{1+r-d+p_2^*(m-d)}{u-d}, \quad p_2^* \in [0,1], \quad p_3^* = 1 - p_1^* - p_2^*.$$

35

所以在三叉树模型中，使用风险中性定价方法得到的结果是不唯一的，有无穷多组价格解.

练习 1.5.4 为了确定所有的风险中性概率测度，关于 d，m，u，r 的哪些条件是必需的? ◀

1.5.3 对冲与资产复制

期权经常被银行用来对交易进行风险对冲. 考虑一份基于股价 S_0 的二叉树模型的欧式期权 C_e，引入记号 $S_1(-)$，$S_1(+)$ 和 $C_e(-)$，$C_e(+)$，后面将看到，该定价公式对一般期权也成立. 现考虑这样一个问题：是否一定存在一种投资组合可以对冲期权交易的风险? 假设存在这样一份复制了期权的资产组合，则可由这份组合来对冲出售期权带来的金融风险. 记银行持有的资产组合为(θ_0, θ_1)，其中 θ_0 表示存入银行账户的货币金额，θ_1 表示持有的股票份额. 设该组合在 t 时刻的价值为 V_t. 如果组合在 $t=0$ 和 $t=1$ 时有相同的价值，则它对冲了期权的风险. 因为

$$V_0 = \theta_0 + \theta_1 S_0,$$

且

$$V_1(\omega) = \begin{cases} \theta_0(1+r) + \theta_1 S_0 u, & \omega = +, \\ \theta_0(1+r) + \theta_1 S_0 d, & \omega = -. \end{cases}$$

而期权在时刻 0 的价值 W_0 为该期权的价格 $\pi(C_e)$，在时刻 1 的价格为

$$W_1(\omega) = \begin{cases} S_0 u - K, & \omega = +, \\ 0, & \omega = -. \end{cases}$$

若对任意的 $\omega \in \Omega$ 和 $t \in \{0, 1\}$，均有 $V_t(\omega) = W_t(\omega)$，那么该组合复制了期权. 于是得到

$$V_0 = \pi(C_e) \tag{1.9}$$

及

$$\theta_0(1+r) + \theta_1 S_0 u = S_0 u - K, \tag{1.10}$$

$$\theta_0(1+r) + \theta_1 S_0 d = 0. \tag{1.11}$$

将 $\theta_0(1+r) = -\theta_1 S_0 d$((1.11)式)代入(1.10)式得到

$$-\theta_1 S_0 d + \theta_1 S_0 u = S_0 u - K \Leftrightarrow \theta_1(S_0 u - S_0 d) = S_0 u - K \tag{1.12}$$

$$\Leftrightarrow \theta_1(S_0 u - S_0 d) = C_e(+) - C_e(-). \tag{1.13}$$

注意到 $S_0 u - S_0 d = S_1(+) - S_1(-)$，所以

$$\theta_1 = \frac{C_e(+) - C_e(-)}{S_1(+) - S_1(-)}.$$

36

该比值(它是复制期权过程中所需要持有的股票份额)称为**对冲比率**. 例 1.1.4 中的对冲比率为 $\theta_1 = 10/20 = 1/2$. 即石油交易商在 $t=0$ 时买入合约数的一半石油来构造对冲组合. θ_0

的计算公式为：

$$\theta_0 = C_e(-) - \frac{C_e(+) - C_e(-)}{u-d} \frac{d}{1+r}.$$

在本例中，$\theta_0 = 0 - \frac{10}{1.1-0.9} \times \frac{0.9}{1.01} \approx -44.55$. 它表明：交易商从银行借入 44.554，同时他收到溢价(或升水)5.45，用这些钱他能买入合约数的一半石油来对冲期权. 对冲的初始成本为 $V_0 = \theta_0 + \theta_1 S_0$，称为**复制成本**，复制成本就是期权的公平价格.

练习 1.5.5 P^* 为由 $p^* = \frac{1+r-d}{u-d}$ 确定的概率测度，求证：$V_0 = E^*\left(\frac{C_e}{1+r}\right)$. ◀

1.5.4 风险中性测度的不存在性

考虑一个有两支满足二叉树模型的股票的金融市场，设股价上升因子分别为 u_1，u_2，下降因子分别为 d_1，d_2. 由风险中性定价得到下列两个方程：

$$p^* u_1 + (1-p^*)d_1 = 1+r,$$
$$p^* u_2 + (1-p^*)d_2 = 1+r,$$

其中 p^* 为自由参数. 由于 p^* 依赖于参数 r，u_1，u_2，d_1，d_2，该方程组可能无解. 因此，风险中性概率测度可能不存在.

1.5.5 Black-Scholes 定价公式

本节将介绍期权定价中著名的 Black-Scholes 公式，其中有些结果只能在以后的章节给出.

假设存在风险中性定价测度 P^*，考虑一份标的资产是股价为 S_t 的股票且敲定价格为 K 的欧式看涨期权，在到期日 T，期权的价值 $C = \max(0, S_T - K)$. 假设利息为常利率，其贴现因子为 e^{-r}，其中 $r > 0$ 为某个常数. 则期权的贴现价格为 $C^* = e^{-rT}\max(0, S_T - K)$. 在风险中性条件下，则有 $E^*(C) = C_0$，其中 C_0 表示随机支付 C 在 $t=0$ 时的公平价格，等价地，有

$$C_0 = E^*(C^*) = e^{-rT}E^*(\max(S_T - K, 0)).$$

这意味着，如果通过上式右边的表达式求出欧式看涨期权的公平价格，前提是必须知道 S_T 在概率测度 P^* 下的分布.

著名的 Black-Scholes 模型假定：在实际概率测度下，股票的对数价格服从漂移为 $\mu \in \mathbf{R}$ 并且波动为 $\sigma > 0$ 的正态分布. 则在 P^* 下，在到期日 T，$\log S_T$ 服从

37

$$\text{漂移为 } \log S + (r - \sigma^2/2)T, \quad \text{波动为 } \sigma\sqrt{T}$$

的正态分布，其中 $S = S_0$ 为股票的当前价格，它是确定期权公平价格的基础.

对数正态分布有下列重要结论：设 X 服从参数为 $m \in \mathbf{R}$，$s > 0$ 的对数正态分布，则有

$$E(X-K)^+ = e^{m+s^2/2}\Phi\left(\frac{m - \log K}{s} + s\right) - K\Phi\left(\frac{m - \log K}{s}\right). \tag{1.14}$$

下面只给出证明的一个框架，具体细节请读者自行完善. 事实上，对 $x \geqslant 0$，有$(X-K)^+ \geqslant$

$x \Leftrightarrow X-K \geqslant x$. 记 X 的密度函数为 $f(x)$，则有

$$E(X-K)^+ = \int_0^\infty P(X \geqslant K+x)\mathrm{d}x$$
$$= \int_0^\infty \int_{K+x}^\infty f(t)\mathrm{d}t\mathrm{d}x$$
$$= \int_K^\infty \int_x^\infty f(t)\mathrm{d}t\mathrm{d}z,$$

其中用到了变量代换 $z=x+K$. 将对数正态分布的密度函数(1.3)式代入，化简后得到

$$E(X-K)^+ = \int_K^\infty \int_x^\infty \frac{1}{\sqrt{2\pi}st} \mathrm{e}^{-(\log t - m)^2/2s^2}\mathrm{d}t\mathrm{d}x.$$

令 $z=(\log t - m)/s$，则 $\mathrm{d}z = \mathrm{d}t/st$，积分化简为

$$\int_K^\infty \int_{(\log x - m)/s}^\infty \varphi(z)\mathrm{d}z\mathrm{d}x,$$

其中 $\varphi(x) = 1/\sqrt{2\pi}\mathrm{e}^{-x^2/2}$ 为标准正态分布的密度函数. 在分布积分公式 $\int uv' = uv| - \int u'v$ 中取 $u(x) = \int_{(\log x - m)/s}^\infty \varphi(z)\mathrm{d}z$，$u'(x)=1$，则有

$$E(X-K)^+ = \int_K^\infty \varphi\left(\frac{\log x - m}{s}\right)\mathrm{d}x - K\Phi\left(\frac{m - \log K}{s}\right),$$

其中，$\Phi(x) = \int_{-\infty}^x \varphi(t)\mathrm{d}t$ 为标准正态分布的分布函数. 最后，作变量代换 $z=\log x$ 则得到式(1.14).

在式(1.14)中取 $m=\log S_0 + (r-\sigma^2/2)T$，$s=\sigma\sqrt{T}$，则得到

$$E^*(S_T - K)^+ = \mathrm{e}^{\log S + rT}\Phi\left(\frac{\log(S/K) + (r-\sigma^2/2)T}{\sigma\sqrt{T}} + \sigma\sqrt{T}\right)$$
$$- K\Phi\left(\frac{\log(S/K) + (r-\sigma^2/2)T}{\sigma\sqrt{T}}\right).$$

因此，欧式看涨期权的公平价格为

$$\pi(C_e) = E^*(C^*) = \mathrm{e}^{-rT}E^*(S_T - K)^+ = S_0\Phi(d_1) - K\Phi(d_2)\mathrm{e}^{-rT}, \quad (1.15)$$

38

其中，

$$d_1 = \frac{\log(S/K) + (r+\sigma^2/2)T}{\sigma\sqrt{T}},$$
$$d_2 = \frac{\log(S/K) + (r-\sigma^2/2)T}{\sigma\sqrt{T}}.$$

利用已经得到的欧式看涨期权的价格，结合看涨看跌平价关系，得到欧式看跌期权的价格为

$$\pi(C_p) = \pi(C_e) + K\mathrm{e}^{-rT} - S.$$

进一步，如果要求期权在 t 时刻的价格，此时期权的到期期限为 $\tau = T-t$，只需把 T 换为 τ，再将 S 换为 t 时刻的标的资产价格即可.

1.5.6 一些希腊字母表示的量

Black-Scholes 定价公式给出了欧式看涨期权的无套利公平价格依赖于哪些变量：除了合约中规定的期权参数 K，T 和初始股价 S_0 外，公式还依赖于无风险利率 r 以及股票对数价格的波动率 σ. 在风险管理中，知道头寸关于这些量的敏感度也是非常必要的. 例如，标的资产的波动率上升，很快会引起期权的价值波动.

1.5.6.1 表示期权一阶导数的希腊字母

本小节给出欧式看涨期权的 Black-Scholes 定价公式的一阶导数(即期权价格对各个变量的偏导数)对应的希腊字母. 这些记号对其他衍生产品同样适用.

为了便于理解，定义期权价格关于股价的敏感度为比率 $\frac{\Delta\pi(C_e)}{\Delta S}$，它表示股价变动一个单位引起的期权价格 V 的变化. 显式期权公式 $\pi(C_e)$ 关于变量 S，T，σ 和 r 是可微函数. 下面给出期权价格关于股价的敏感度的严格数学定义，称为 **Delta 系数**，定义：

$$\Delta=\frac{\partial\pi(C_e)}{\partial S}.$$

同样，我们可以定义期权价格关于到期日 T 变动的敏感度，称为 **Theta 系数**，它为

$$\Theta=\frac{\partial\pi(C_e)}{\partial T}.$$

而 **Vega**(或 **Kappa**)**系数**是期权价格关于波动率的敏感度，它为

39

$$\nu=\frac{\partial\pi(C_e)}{\partial\sigma}.$$

最后，**Rho 系数**表示期权价格关于利率变动的敏感度，表达式为

$$\rho=\frac{\partial\pi(C_e)}{\partial r}.$$

表 1-2 给出了 Black-Scholes 模型对应的希腊字母的表达式.

表 1-2 欧式期权的希腊字母

希腊字母	看涨期权	看跌期权
$\Delta=\frac{\partial\pi(C_e)}{\partial S}$	$\Phi(d_1)$	$\Phi(d_1)-1$
$\Theta=\frac{\partial\pi(C_e)}{\partial T}$	$-\frac{S_\varphi(d_1)\sigma}{2\sqrt{T}}-rK\mathrm{e}^{-rT}\Phi(d_2)$	$-\frac{S_\varphi(d_1)\sigma}{2\sqrt{T}}+rK\mathrm{e}^{-rT}\Phi(-d_2)$
$\nu=\frac{\partial\pi(C_e)}{\partial\sigma}$	$S\varphi(d_1)\sqrt{T}$	$S\varphi(d_1)\sqrt{T}$
$\rho=\frac{\partial\pi(C_e)}{\partial r}$	$KT\mathrm{e}^{-rT}\Phi(d_2)$	$-KT\mathrm{e}^{-rT}\Phi(-d_2)$
$\Gamma=\frac{\partial^2\pi(C_e)}{\partial S^2}$	$\frac{\varphi(d_1)}{S\sigma\sqrt{T}}$	$\frac{\varphi(d_1)}{S\sigma\sqrt{T}}$

利用表示一阶导数的希腊字母，可以将期权价格近似表示为一个线性函数. 例如，如标的资产的价格从 S 变化到 $\widetilde{S}$，且已知 $\Delta=\frac{\partial\pi(C_e)}{\partial S}$，则可推出

$$\pi \approx \pi(C_e) + \frac{\partial \pi(C_e)}{\partial S}(\widetilde{S} - S),$$

当$|\widetilde{S}-S|$较小时，可用此近似结果.

需要指出的是，上面这些一阶偏导数仍然是其余参数的函数，因此，它们的值仍然依赖于这些参数的取值. 如果同时有几个参数发生变化，那么，从单个的敏感度就无法了解期权价格的反馈.

给定了期权敲定价 K 及 S，T，σ 和 r，则唯一地确定了 $\pi(C_e)$. 显然，上述几个希腊字母构成一个梯度，即

$$\frac{\partial \pi(C_e)}{\partial \vartheta} = \left(\frac{\partial \pi(C_e)}{\partial S}, \frac{\partial \pi(C_e)}{\partial T}, \frac{\partial \pi(C_e)}{\partial \sigma}, \frac{\partial \pi(C_e)}{\partial r}\right)' = (\Delta, \Theta, \nu, \rho)',$$

其中，$\vartheta=(S, T, \sigma, r)'$. 当参数由 ϑ 变化为 $\widetilde{\vartheta}=(\widetilde{S}, \widetilde{T}, \widetilde{\sigma}, \widetilde{r})'$时，由泰勒公式得到期权的线性逼近公式为

$$\pi \approx \pi(C_e) + \frac{\partial \pi(C_e)}{\partial \vartheta}(\widetilde{\vartheta} - \vartheta) = \pi(C_e) + \Delta(\widetilde{S} - S) + \Theta(\widetilde{T} - T) + \nu(\widetilde{\sigma} - \sigma) + \rho(\widetilde{r}, - r).$$

40

1.5.6.2 表示二阶偏导数的希腊字母

由一阶偏导数对应的希腊字母可以得到期权的线性逼近公式. 下一步就是考虑二阶偏导数，它可以推导出二次逼近.

人们主要感兴趣的是期权价格对股价变化的依赖程度，于是引出了 **Gamma 系数**的定义，它为

$$\Gamma = \frac{\partial^2 \pi(C_e)}{\partial S^2}.$$

1.5.7 模型校验方法、隐含波动率和波动率微笑

为了利用 Black-Scholes 公式进行期权定价，必须确定利率 r 的值. 通常它被取为投资者是在短期内收到的投资收益. 除此之外，还需要确定波动率 σ 的值，而 σ 是无法直接观测的. 求 σ 的常见方法有两种，一种是在 1.3.2 节中介绍的统计学方法，它是用历史数据来估计 σ 的值；另外一种是**模型校验**方法，即将金融市场上的实际数据代入公式，从而求出公式中的未定参变量. 该方法有一个优点，它可以重新产生一组数据，从而可以与市场实际数据进行比较，因此很受投资者的欢迎，因为投资者都不会喜欢一个不可信或者与实际市场相矛盾的模型.

利用 Black-Scholes 公式，通过变化自由参数 σ 的值，将模型算出的期权价格和实际市场价格相比较，从而得到 σ 的理想值. 注意，在令方程(1.15)等于实际价格后，得到的是一个 σ 的非线性方程，它是在固定的敲定价格 K 和到期期限 $T-t$ 的条件下进行的. 用这种方法确定的 σ 值称为**隐含波动率**.

理论上，标的资产的波动率 σ 是与敲定价格和到期期限无关的常数. 然而，通过改变 K 和 T 的值而分别求出隐含波动率，可以发现求出的隐含波动率是与这两个参数有关的.

有时，隐含波动率是 K 的递减函数，这种现象称为**波动率偏离**；有时，特别是关于外币交易的期权，价期权的波动率较低，而在期权进入价内或价外状态时波动率上升，这种效应称为**波动率微笑**. 隐含波动率对 K 的依赖性一般用**货币性**或**敲定比**(即 S/K)来反映. 如果在三维空间中表出隐含波动率与敲定价 K(或 K/S)和到期日 T 的关系，则可以得到一个二维曲线，即所谓的**波动率曲面**. 图 1-2 给出的是 SIEMENS AG 的波动率曲面.

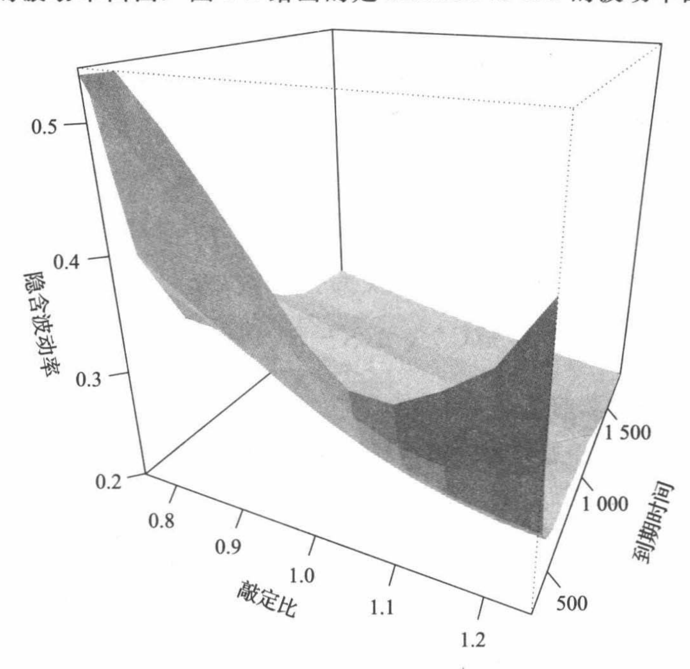


图 1-2 SIEMENS AG 在 11 月 4 日的欧式看涨期权波动率曲面(期权有效期从 2011-11 到 2015-12，到期时间以天为单位. 数据来源于 DATASTREAM)

1.5.8 期权价格与风险中性密度

期权价格与风险中性概率密度之间存在着有趣而且重要的关系. 这种关系并不是只对 Black-Scholes 模型成立，它是金融市场固有的性质. 利用这种关系可以从期权价格推出风险中性概率.

41

在前面推导 Black-Scholes 公式时，得到了下列的公式

$$C_e(K) = \mathrm{e}^{-rT}E^*(S_T - K)^+, \tag{1.16}$$

现对它稍作变动，把它看作敲定价格 K 的函数. 把风险中性价格仍记为 $C_e(K)$，但撇开了 Black-Scholes 模型. 假设存在一个风险中性概率测度 P^* 可以用来定价未来的期权价格，同时假设期权到期时股价 S_T 在概率测度 P^* 下的概率密度函数为 $\varphi_T^*(x)$，则有

$$P^*(S_T \leqslant x) = \int_{-\infty}^{x} \varphi_T^*(u)\mathrm{d}u, \quad x \in \mathbf{R}.$$

于是，(1.16)式可以改写为

$$C_e(K) = \mathrm{e}^{-rT}\int_{-\infty}^{\infty}(x-K)^+ \varphi_T^*(x)\mathrm{d}x = \mathrm{e}^{-rT}\int_K^{\infty}\varphi_T^*(x)\mathrm{d}x.$$

注意到

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_{a(t)}^{b(t)} f(x,t)\mathrm{d}x = f(b(t),t)b'(t) - f(a(t),t)a'(t) + \int_{a(t)}^{b(t)}\frac{\partial f(x,t)}{\partial t}\mathrm{d}x,$$

42

从而得到欧式看涨期权风险中性价格的一阶导数(关于 K)为

$$\frac{\partial C_e(K)}{\partial K} = -\int_K^{\infty}\varphi_T^*(x)\mathrm{d}x,$$

二阶导数为

$$\frac{\partial^2 C_e(K)}{\partial K^2} = \varphi_T^*(K).$$

因此，通过分析不同敲定价格 K 的期权价格可以确定风险中性概率测度. 同时由任意的密度函数都是非负的可知，期权价格是关于敲定价格的凸函数.

1.6 评注与延伸阅读

一本不涉及复杂数学推导的关于期权、远期和约及衍生产品的综述材料是畅销书 Hull(2009)，它给出了许多生动的例子，详尽地解释隐含在金融工具中的经济学原理和金融市场的运作方式，并给出了金融工具定价的基本公式. 对其中的数学理论，它以一种通俗易懂的方式进行处理. 关于介绍数理金融的书籍，我们也推荐 Baird(1992)、Pliska(1997)和 Buchanan(2006)，他们主要介绍离散时间及有限概率空间中的模型. 关于风险测度理论，我们要感谢 Artzner 等(1999)的启发性工作，以及 Pflug 和 Römisch(2007)与 Embrechts 等(2002)最近的专著，他们给出了有关测度的许多性质. 关于统计学的基本理论，如估计、假设检验和置信区间等，有许多文献可供参考，比如，Lehmann 和 Romano(2005)，或 Shao(2003). Lai 和 Xing(2008)介绍了金融统计. 关于核光滑方法及其性质的更多知识可参考 Silverman(1986)，Härdle(1990)，Fan 和 Gijbels(1996)以及 Wand 和 Jones(1995). Scott 和 Terrell(1987)，以及 Savchuk 等(2010)介绍了如何选取最优带宽. 最近 Golyandina 等(2011)使用奇异谱分析方法讨论了如何选取最优带宽.

参考文献

Artzner P., Delbaen F., Eber J.M. and Heath D. (1999) Coherent measures of risk. *Math. Finance* **9**(3), 203–228.

Bachelier L. (1900) Théorie de la spéculation. *Annales Scientifiques de l' É.N.S.* 3e **17**, 21–86.

Baird A.J. (1992) *Option Market Making: Trading and Risk Analysis for the Financial and Commodity Option Markets*. Wiley Finance, Wiley & Sons.

Buchanan J. (2006) *An Undergraduate Introduction to Mathematical Finance*. World Scientific, Singapore.

Embrechts P., McNeil A.J. and Straumann D. (2002) Correlation and dependence in risk management: properties and pitfalls *Risk Management: Value at Risk and Beyond (Cambridge, 1998)*. Cambridge Univ. Press Cambridge pp. 176–223.

43

Fan J. and Gijbels I. (1996) *Local Polynomial Modelling and its Applications*. vol. 66 of *Monographs on Statistics and Applied Probability*. Chapman & Hall, London.

Golyandina N., Pepelyshev A. and Steland A. (2011) New approaches to nonparametric density estimation and selection of smoothing parameters. **56**(7), 2206–2218.

Härdle W. (1990) *Applied Nonparametric Regression*. vol. 19 of *Econometric Society Monographs*. Cambridge University Press, Cambridge.

Hull J.C. (2009) *Options, Futures, and Other Derivatives*. 7th edn. Pearson Prentice Hall.

Lai T.L. and Xing H. (2008) *Statistical Models and Methods for Financial Markets*. Springer Texts in Statistics. Springer, New York.

Lehmann E.L. and Romano J.P. (2005) *Testing Statistical Hypotheses*. Springer Texts in Statistics third edn. Springer, New York.

Pflug G.C. and Römisch W. (2007) *Modeling, Measuring and Managing Risk*. World Scientific Publishing Co. Pte. Ltd., Hackensack, NJ.

Pliska S. (1997) *Introduction to Mathematical Finance*. Blackwell Publishing, Oxford.

Savchuk O.Y., Hart J.D. and Sheather S.J. (2010) Indirect cross-validation for density estimation. *J. Amer. Statist. Assoc.* **105**(489), 415–423. With supplementary material available online.

Scott D.W. and Terrell G.R. (1987) Biased and unbiased cross-validation in density estimation. *J. Amer. Statist. Assoc.* **82**(400), 1131–1146.

Shao J. (2003) *Mathematical Statistics*. Springer Texts in Statistics 2nd edn. Springer-Verlag, New York.

Silverman B.W. (1986) *Density Estimation for Statistics and Data Analysis*. Monographs on Statistics and Applied Probability. Chapman & Hall, London.

Wand M.P. and Jones M.C. (1995) *Kernel Smoothing*. vol. 60 of *Monographs on Statistics and Applied Probability*. Chapman and Hall Ltd., London.

第 2 章　单期模型的套利理论

本章对期权定价作进一步的研究，主要讨论未定权益为任意随机支付的情形. 为了便于快速入门，本章只讨论单期模型. 为了简化证明过程，有些结论是建立在有限样本空间上的，即结论是在假定只有有限的样本个数条件下证明的.

上一章用风险中性定价原理已经能够得到期权价格的一个简单而且漂亮的结果，本章将给出定价测度存在及唯一的条件，并讨论该测度的计算方法. 当定价测度存在时，由它得到的期权价格是否真的排除了套利机会？另外一个有趣的问题是：未定权益风险在多大程度上能够被对冲？或者说，能够复制一个与它完全等价的资产组合吗？这些问题都将在本章被讨论.

2.1　定义与预备

考虑一个金融市场，它有一个无风险投资(如国库券或银行存款)和 d 种风险投资. 风险投资通常是指投资股票，本书所涉及的风险资产专指股票资产，但是本书所用的模型同样适用于其他投资，因为对风险资产的唯一要求是其价格必须是随机的. 单期模型是假设投资者只能在 $t=0$ 和 $t=1$ 两个时刻交易. 在 $t=0$ 时刻投资者建立资产组合，而在 $t=1$ 时刻资产价值为其清仓所得.

在详细介绍模型之前，首先引入一些记号：记 r 为债券或银行的无风险固定利率，且假设存贷利率均为 r. 用 x 表示 $t=0$ 时的初始资产，对 x 的符号规定如下：若 $x>0$，表示拥有一笔存款；若 $x<0$，表示投资者产生了一项额度为 $|x|$ 的信用贷款. 在 $t=1$ 时，该款项的价值为

$$v=(1+r)x.$$

同样，对无风险债券投资，如债券初始价格为 $S_{00}\in(0,\infty)$，则在 $t=1$ 时，它的价值为

$$S_{10}=(1+r)S_{00}.$$

为了简化符号，以后假定 $S_{00}=1$.

分别用 S_{01}，…，S_{0d} 表示 d 个风险资产的价格. 记基本概率空间为 $(\Omega,\mathcal{F},P)$，每个样本点 $\omega\in\Omega$ 对应 $t=1$ 时市场可能出现的一个结果. P 为实际概率测度，即对每个可测集 $A\in\mathcal{F}$，其发生概率为 $P(A)$.

$t=1$ 时刻市场的未知风险价格可以用下列随机变量来表示

$$S_{10},\cdots,S_{1d}:(\Omega,\mathcal{F},P)\to(\mathbf{R},\mathcal{B})$$

其中 $S_{10}=(1+r)S_{00}$. 称随机向量 $S_0=(S_{00},\cdots,S_{0d})\in\mathbf{R}^{d+1}$ 为**价格向量**.

定义 2.1.1　(ⅰ) 设 $S_0=(S_{00},\cdots,S_{0d})'$ 为一个已知的价格向量，即对 $i=0$，…，d，$S_{0i}>0$ 是已知的初始价格. 又设 $S_1=(S_{10},\cdots,S_{1d})'$ 为一个满足 $S_{1i}:(\Omega,\mathcal{F},P)\to(\mathbf{R},\mathcal{B})$，$i=1$，…，$d$ 均为正随机变量的随机向量，则称 $\{S_t:t=0,1\}$ 为一个**价格过程**.

(ⅱ) $\{S_t^*:t=0,1\}$ 称为**贴现价格过程**，如果它满足 $S_0^*=S_0$ 且 $S_1^*=(S_{10}^*,\cdots,S_{1d}^*)'$，

其中，$S_{1i}^*=\dfrac{S_{1i}}{1+r}$，$i=0, 1, \cdots, d$.

有了上述定义，就可以从数学意义上把金融市场定义为$((\Omega, \mathcal{F}, P), \{S_t\})$.

定义 2.1.2(资产组合) 称向量 $\varphi=(\varphi_0, \cdots, \varphi_d)'\in\mathbf{R}^{d+1}$ 为**资产组合**，称它的分向量$(\varphi_1, \cdots, \varphi_d)'$为**风险资产组合**.

规定：$\varphi_i>0$ 表示持有资产多头头寸，$\varphi_i<0$ 表示持有资产空头头寸.

定义 2.1.3 (i) 对任意一个资产组合 $\varphi=(\varphi_0, \cdots, \varphi_d)'\in\mathbf{R}^{d+1}$和价格过程$\{S_t: t=0, 1\}$，称$\{V_t: t=0, 1\}$为**资产组合 $\boldsymbol{\varphi}$ 的价值过程**，其中，

46

$$V_t = V_t(\varphi) = \varphi' S_t = \sum_{i=0}^{d} \varphi_i S_{ti}, \quad t = 0,1.$$

(ii) **贴现价值过程**$\{V_t^*: t=0, 1\}$的定义如下：

$$V_t^* = V_t^*(\varphi) = \varphi' S_t^*, \quad t = 1.$$

显然，V_t 表示资产组合 φ 的价值，要求 $\varphi_i\geqslant 0$. 因为购买 φ_i 份资产 i 需要支付的本金为 $\varphi_i S_{0i}$，则资产组合的初始资本为：

$$V_0 = \varphi_0 + \sum_{i=1}^{d} \varphi_i S_{0i}.$$

其中，$\sum_{i=1}^{d}\varphi_i S_{0i}$ 用来购买风险资产，若为正，则 $V_0-\sum_{i=1}^{d}\varphi_i S_{0i}$ 为存入银行账户金额；若为负，则需从银行借入 $V_0-\sum_{i=1}^{d}\varphi_i S_{0i}$ 金额用来建仓. 资产组合里的每一项初始空头头寸，即 $\varphi_i<0$ 的资产，都可消耗掉$|\varphi_i S_{0i}|$的初始多头头寸.

定义 2.1.4 (i) 称$\{G_t: t=0, 1\}$为**收益过程**，其中，$G_t=G_t(\varphi)=V_t(\varphi)-V_0(\varphi)$，$t=0, 1$.

(ii) 称$\{G_t^*: t=0, 1\}$称为**贴现收益过程**，其中，$G_t^*=G_t(\varphi)=V_t^*(\varphi)-V_0^*(\varphi)$，$t=0, 1$.

2.2 线性定价测度

本节研究在什么条件下资产组合能带来无风险的利润，令人惊讶的结果是：它与风险中性测度的存在性密切相关，而风险中性测度是使用风险中性原理给未定权益定价的有力工具. 为了得到这种对应关系，下面引入占优资产组合定义，它比套利更强，且更受到投资者欢迎.

定义 2.2.1 一个资产组合 φ 称为**占优的**，如果存在另外一个资产组合$\widetilde{\varphi}\neq\varphi$，有 $V_0(\varphi)=V_0(\widetilde{\varphi})$且 $V_1(\varphi)>V_1(\widetilde{\varphi})$($P$-a. s.)成立. 亦称 φ 优于$\widetilde{\varphi}$.

引理 2.2.2 占优资产组合存在的充分必要条件是存在某个资产组合 φ，使得

$$V_0(\varphi) = 0 \quad 且 \quad V_1(\varphi) > 0 \quad P\text{-a. s.}$$

证明

"$\Rightarrow$"：设 $\overline{\varphi}$ 优于$\widetilde{\varphi}$，令 $\varphi=\overline{\varphi}-\widetilde{\varphi}$，由映射 $\varphi\mapsto V(\varphi)$的线性性质知

$$V_0(\varphi) = V_0(\overline{\varphi}) - V_0(\widetilde{\varphi}) = 0, \quad 且\ V_1(\varphi) = V_1(\overline{\varphi}) - V_1(\widetilde{\varphi}) > 0(\text{a. s.})\ 成立.$$

"$\Leftarrow$"：显然，满足 $V_0(\varphi)=0$ 且 $V_1(\varphi)>0$(a.s.)的投资比不投资好. ■ 47

命题 2.2.3 设 Ω 为一个有限集，则存在占优资产组合的充分必要条件是存在某个资产组合 $\widetilde{\varphi}$，使得

$$V_0(\widetilde{\varphi})<0 \quad 且 \quad V_1(\widetilde{\varphi})\geqslant 0 \quad P\text{-a.s.}$$

证明 只证必要性. 由引理 2.2.2 知，存在满足 $V_0^*(\varphi)=0$ 且 $V_1^*(\varphi)>0$(P-a.s.)的资产组合 φ，从而有

$$G_1^*(\varphi)(\omega)=V_1^*(\varphi)(\omega)-V_0(\varphi)>0$$

对所有 $\omega\in\Omega$ 成立. 因此

$$\delta:=\min_{\omega\in\Omega}G_1^*(\varphi)(\omega)>0.$$

定义 $\widetilde{\varphi}\in\mathbf{R}^{d+1}$ 为 $\widetilde{\varphi}_i=\varphi_i$，$i=1$，…，$d$；$\widetilde{\varphi}_0=-\sum_{i=1}^{d}\varphi_iS_{0i}^*-\delta$. 则得到

$$V_0^*(\widetilde{\varphi})=\widetilde{\varphi}_0+\sum_{i=1}^{d}\widetilde{\varphi}_iS_{0i}^*=-\delta<0.$$

且对任意 $\omega=\Omega$，有

$$\begin{aligned}V_1^*(\widetilde{\varphi})(\omega)&=\widetilde{\varphi}_0+\sum_{i=1}^{d}\varphi_iS_{1i}^*(\omega)\\&=-\sum_{i=1}^{d}\varphi_iS_{0i}^*(\omega)-\delta+\sum_{i=1}^{d}\varphi_iS_{1i}^*(\omega)\\&=-\delta+\sum_{i=1}^{d}\varphi_i(S_{1i}^*-S_{0i})(\omega)\\&=-\delta+G_1^*(\widetilde{\varphi})(\omega)\geqslant 0.\end{aligned}$$ ■

由定义知：存在占优资产组合意味着存在两个具有同样价格的资产组合，其中一个的收益几乎必然(以概率 1)比另外一个好. 当然，如果 $t=1$ 时的价格是成本的单调函数，这种情况不会发生.

定义 2.2.4 (i) 设 Ω 为一个有限集，称 $\pi=\{\pi(\omega):\omega\in\Omega\}\subset[0,\infty)$ 为一个(**线性**)**定价测度**，如果对任意的 $\varphi\in\mathbf{R}^{d+1}$，均有

$$V_0^*(\varphi)=\sum_{\omega\in\Omega}\pi(\omega)V_1^*(\omega).$$

48

(ii) 对任意 Ω，设$\mathcal{F}$是 Ω 中的 σ 代数，称测度 π 为**定价测度**，如果

$$V_0^*(\varphi)=\int V_1^*(\omega)\mathrm{d}\pi(\omega),\quad 对所有\ \varphi\in\mathbf{R}^{d+1}\ 成立.$$

对给定的一个定价测度 π，易证它是一个概率测度，即满足

$$\sum_{\omega\in\Omega}\pi(\omega)=1,\quad 且\int_{\Omega}\mathrm{d}\pi=1.$$

引理 2.2.5 设 Ω 为一个有限集.

(i) 对任意的定价测度 $\{\pi(\omega):\omega\in\Omega\}$，有

$$S_{0j}=\sum_{\omega}\pi(\omega)S_{1j}^*(\omega),\quad j=1,\cdots,d. \tag{2.1}$$

(ii) 如果$\{\pi(\omega):\omega\in\Omega\}$是一个满足(2.1)式的概率测度，那么它是一个定价测度.

证明 (i) 对固定的$j\in\{1,\cdots,d\}$，考虑资产组合φ，其中$\varphi_i=\mathbf{1}(i=j)$，$i=0,\cdots,d$，则$V_0(\varphi)=S_{0j}$，$V_1^*(\varphi)=S_{1j}^*$. 如果$\{\pi(\omega)\}$是一个定价测度，则

$$S_{0j}=V_0(\varphi)=\sum_{\omega\in\Omega}\pi(\omega)V_1^*(\omega)=\sum_{\omega\in\Omega}\pi(\omega)S_{1j}^*(\omega).$$

(ii) 结论由$\varphi\mapsto V_t(\varphi)$是一个线性映射马上得到. ■

由引理2.2.5知，初始价格是未来的可能价格被定价测度加权后的线性函数. 其线性结构使得人们能够用线性方程组$\boldsymbol{Ax}=\boldsymbol{b}$来处理定价问题，其中$\boldsymbol{A}$是$n\times p$的矩阵，$\boldsymbol{b}$是$n$维向量，$\boldsymbol{x}$是$p$维未知解向量. 具体方法如下：若$\Omega=\{\omega_1,\cdots,\omega_n\}$是有限的，$\pi=\{\pi(\omega)\}$是一个定价测度，记$\pi_i=\pi(\omega_i)$，$i=1,\cdots,n$，则有

$$S_{0j}^*=\sum_{i=1}^{n}\pi_i\cdot S_{1j}^*(\omega_i),\quad j=1,\cdots,d.$$

反之，上述方程组的任意一个满足条件

49

$$0\leqslant\pi_1,\cdots,\pi_n\leqslant 1,\quad \sum_{i=1}^{n}\pi_i=1$$

的解向量$(\pi_1,\cdots,\pi_n)'$构成一个定价测度. 使用矩阵记号，则有

$$\boldsymbol{D}^*\pi=\boldsymbol{b}^*,$$

其中

$$\boldsymbol{D}^*=\begin{pmatrix}1&\cdots&1\\S_{11}^*(\omega_1)&\cdots&S_{11}^*(\omega_n)\\\vdots&\vdots&\vdots\\S_{1d}^*(\omega_1)&\cdots&S_{1d}^*(\omega_n)\end{pmatrix},\quad \boldsymbol{b}^*=\begin{pmatrix}1\\S_{01}^*\\\vdots\\S_{0d}^*\end{pmatrix}$$

称$\boldsymbol{D}^*$为**贴现支付矩阵**或**贴现收益矩阵**.

下列定理说明了定价测度与占优资产组合之间的关系.

定理2.2.6 设$\Omega=\{\omega_1,\cdots,\omega_n\}$是有限集，存在线性定价测度的充分必要条件是不存在占优资产组合.

证明 因为求解$\pi\geqslant 0$，$\boldsymbol{D}^*\pi=\boldsymbol{b}^*$等价于求解线性规划问题

$$\max_{\pi\in\mathbf{R}^n}\mathbf{0}'\pi \text{ 使得 } \pi\geqslant 0,\boldsymbol{D}^*\pi=\boldsymbol{b}^*,\tag{P}$$

其中$\mathbf{0}\in\mathbf{R}^n$表示零向量，$\boldsymbol{D}^*$为贴现支付矩阵. 如果存在定价测度$\pi$，则问题$(P)$有解，由线性规划的对偶理论知，它等价于下列对偶问题

$$\min_{\varphi\in\mathbf{R}^{d+1}}\varphi'\boldsymbol{b}^* \text{ 使得 } \varphi'\boldsymbol{D}^*\geqslant 0\tag{D}$$

存在最优解，且它们的最优值一致. 而问题(D)的解φ是满足$V_0(\varphi)=V_0^*(\varphi)=\varphi'\boldsymbol{b}^*=0$，及$((\boldsymbol{D}^*)'\varphi)_k=\varphi_0+\sum_{i=1}^{d}\varphi_iS_{1i}^*(\omega_k)=V_1^*(\varphi)(\omega_k)\geqslant 0$，$k=1,\cdots,n$的资产组合，即$\varphi$是使价格$V_0(\varphi)=0$且$V_1^*(\varphi)>0$极小化的投资组合. 由命题2.2.3知，不存在投资组合使得$V_0<0$，$V_1\geqslant 0$，必要性得证.

为证充分性，假设不存在占优资产组合，则由命题 2.2.3 知，不存在 $V_0(\varphi)<0$ 且对所有 $\omega\in\Omega$ 均有 $V_1(\varphi)(\omega)\geqslant 0$ 的资产组合 φ. 因为满足 $V_1\geqslant 0$ 的资产组合必须满足 $V_0\geqslant 0$，所以 $\varphi=\mathbf{0}\in\mathbf{R}^{d+1}$满足 $V_0(\varphi)=V_1(\varphi)=0$，且是$(D)$的解. $\varphi=0$ 对应的(P)问题的解就是所求的定价测度. ■

2.3 套利理论的进一步讨论

在定义 1.4.1 中，已经引入了套利机会的定义. 本节把它引入具体的模型中，并介绍它的一些等价性质.

50

定义 2.3.1 称资产组合 $\varphi\in\mathbf{R}^{d+1}$ 为**套利机会**或**套利资产组合**，如果它满足 $\varphi' S_0\leqslant 0$(无初始成本)，且有

$$\varphi' S_1 \geqslant 0(P\text{-a. s.})\ 且\ P(\varphi' S_1 > 0) > 0.$$

该定义意味着，在时刻 1 的资产收益几乎必然是非负的，并且获得实际正收益的概率大于零.

注 2.3.2 资产组合的套利性质与实际概率测度 P 有关，但套利在下列集合上是不变的

$$\mathcal{P} = \mathcal{P}(P) = \{\widetilde{P}:\widetilde{P}是(\Omega,\mathcal{F})\ 上的概率测度,且\widetilde{P}\sim P\}.$$

证明留给读者.

回忆一下下列概率论概念：$(\Omega,\ \mathcal{F})$上的两个概率测度 ν 和 μ 称为**等价的**，记为 $\nu\sim\mu$，如果它们有共同的零测集，即对任意的 $A\in\mathcal{F}$，均有

$$\mu(A) = 0 \quad \Leftrightarrow \quad \nu(A) = 0.$$

定义 2.3.3 称一个金融市场是**无套利的**，如果该市场不存在套利机会. 此时，市场满足**无套利条件**.

因为没有人确切地知道未来会发生什么情况，因此无套利性质仅仅是对金融市场数学模型的一个理论假设. 该模型定义由可测空间$(\Omega,\ \mathcal{F})$、价格过程 S_t 和概率测度 P 共同构成.

无套利条件也适用于 d 维收益过程的情形.

引理 2.3.4 金融市场无套利的充分必要条件是对任意的股票资产组合 $\varphi=(\varphi_1,\ \cdots,\ \varphi_d)'$，下列结论均成立

$$\varphi' G_1^* \geqslant 0(\text{a. s.}) \quad \Rightarrow \quad \varphi' G_1^* = 0(\text{a. s.})$$

证明 "$\Rightarrow$"：设市场是无套利的，对满足 $\varphi' G_1^* > 0(\text{a. s.})$的资产组合$(\varphi_1,\ \cdots,\ \varphi_d)\in\mathbf{R}^d$，如果 $\varphi' G_1^*>0$ 有正概率. 下面将证明存在某个 $\widetilde{\varphi}_0\in\mathbf{R}$，使得 $\widetilde{\varphi}=(\widetilde{\varphi}_0,\ \varphi_1,\ \cdots,\ \varphi_d)'$是一个有套利的资产组合，即

51

$$\widetilde{\varphi}' S_0 \leqslant 0, \quad \widetilde{\varphi}' S_1 \geqslant 0(\text{a. s.}), \quad P(\widetilde{\varphi}' S_1 > 0) > 0.$$

事实上，由假定知

$$\varphi' G_1^* = \sum_{i=1}^d \varphi_i(S_{1i}^* - S_{0i}) = \sum_{i=1}^d \varphi_i S_{1i}^* - \sum_{i=1}^d \varphi_i S_{0i} > 0$$

$$\overset{(1+r)}{\Rightarrow} \qquad \sum_{i=1}^d \varphi_i S_{1i} - \sum_{i=1}^d \varphi_i S_{0i}(1+r) > 0$$

有正概率. 取 $\widetilde{\varphi}_0=-\sum_{i=1}^{d}\varphi_i S_{0i}$，即贷款额为 φ_1，…，φ_d 的资产组合，令 $\widetilde{\varphi}=(\widetilde{\varphi}_0, \varphi_1, \cdots, \varphi_d)$，那么

$$\varphi' S_0 = 0$$

且

$$\varphi' S_1 = \underbrace{-\sum_{i=1}^{d}\varphi_i S_{0i}(1+r)}_{\text{信贷值}} + \sum_{i=1}^{d}\varphi_i S_{1i} > 0$$

有正概率. 因此 $\widetilde{\varphi}$ 是一个有套利的组合，与已知矛盾.

“⇐”：因为对任意的股票组合 $\widetilde{\varphi}=(\widetilde{\varphi}_1, \cdots, \widetilde{\varphi}_d)$，均有

$$\widetilde{\varphi}' G_1^* \geqslant 0 \quad \text{a. s.} \quad \Rightarrow \quad \widetilde{\varphi}' G_1^* = 0 \quad \text{a. s.}, \tag{2.2}$$

假设 $\varphi=(\varphi_0, \varphi_1, \cdots, \varphi_d)$ 为一个有套利的组合，即

$$\varphi' S_0 \leqslant 0, \quad \varphi' S_1 \geqslant 0 \quad \text{a. s.}, P(\varphi' S_1 > 0) > 0. \tag{2.3}$$

则有

$$\begin{aligned}
\varphi' G_1^* &= \sum_{i=1}^{d}\varphi_i S_{1i}^* - \sum_{i=1}^{d}\varphi_i S_{0i} \\
&= \sum_{i=0}^{d}\varphi_i S_{1i}^* - \underbrace{\sum_{i=0}^{d}\varphi_i S_{0i}}_{\leqslant 0(\text{a. s.})} \\
&\geqslant \sum_{i=0}^{d}\varphi_i S_{1i} \cdot \frac{1}{1+r} = A.
\end{aligned}$$

因为 $A=\varphi' S_1/(1+r)$，由(2.3)式知 $A\geqslant 0$(a. e.)且 $P(A>0)>0$，这与(2.2)式矛盾，所以市场上不存在套利组合. ■

2.4 $\mathbf{R}^n$ 空间上的分离定理

本节将讨论空间上的几个重要分离定理，在后面的讨论中将会用到它们. 这些定理依据一个线性函数的值来分割出某个子集，即根据对应的函数值来判断每一个 x 是否属于该子集. 下面引入几个具体的命题及其证明，先从一个最基本的定理开始.

定理 2.4.1 设 $K\subset\mathbf{R}^n$ 是一个闭凸集，且 $0\notin K$. 则存在某个 $\lambda\in\mathbf{R}^n$ 和一个实数 $a>0$，使得

$$\lambda' x \geqslant a, \quad \text{对所有 } x\in K \text{ 成立},$$

其中 $a=\text{dist}(0, K)^2$.

证明 设 $r>0$ 为一个实数，构造一个以 0 为中心，r 为半径的闭球，记为

$$\overline{B(0,r)} = \{y\in\mathbf{R}^n : \|y\|\leqslant r\},$$

取 r 足够大而使之能与集合 K 相交. 记 $M=K\cap\overline{B(0, r)}\neq\varnothing$，易知 M 是紧集. 因此，连续映射

$$x\mapsto\|x\|, \quad x\in\mathbf{R}^n,$$

在 M 上可以取到最小值. 记该唯一的最小值点为 $x_0\in M$，则有

$$\|x\|\geqslant\|x_0\| \quad \text{对所有 } x\in K \text{ 成立.} \tag{2.4}$$

因为 K 是凸集，对任意的 $x\in K$ 和 $\alpha\in[0,1]$，则有

$$\alpha x+(1-\alpha)x_0=x_0+\alpha(x-x_0)\in K.$$

因此，在(2.4)式中令 $x=x_0+\alpha(x-x_0)$，可以得到

$$\|x_0+\alpha(x-x_0)\|\geqslant\|x_0\| \quad \Leftrightarrow \quad (x_0+\alpha(x-x_0))'(x_0+\alpha(x-x_0))\geqslant x_0'x_0.$$

最后一个不等式等价于

$$2\alpha(x-x_0)'x_0+\alpha^2(x-x_0)'(x-x_0)\geqslant 0, \quad \alpha\in[0,1].$$

所以对任意的 $\alpha\to 0$，有 $(x-x_0)'x_0\geqslant 0$. 现令 $\lambda:=x_0$，则有

$$\lambda'x\geqslant a:=\lambda'\lambda, \quad \text{对所有 } x\in K \text{ 成立.}$$ ■

作为准备，下面引入一个简单的结论.

引理 2.4.2 设 $\{\lambda,\lambda_k\}\subset\mathbf{R}^n$ 和 $\{x,x_k\}\subset\mathbf{R}^n$ 是欧式空间中的两个点列，且当 $k\to\infty$ 时，$\lambda_k\to\lambda$，$x_k\to x$，则有

$$\lambda_k'x_k\to\lambda'x, \quad k\to\infty.$$

证明 当 $k\to\infty$ 时，由 $\lambda_k\to\lambda$ 知 $\|\lambda_k\|\to\|\lambda\|$，因此存在某个常数 $C>0$，使得 $\|\lambda_k\|\leqslant C$. 由分解式 $\lambda_k'x_k-\lambda'x=\lambda_k'(x_k-x)+(\lambda_k-\lambda)'x$ 可得

$$\|\lambda_k'x_k-\lambda'x\|\leqslant\|\lambda_k\|\,\|x_k-x\|+\|\lambda_k-\lambda\|\,\|x\|.$$

现很容易得到下列结论. ■

定理 2.4.3 设 $K\subset\mathbf{R}^n$ 是一个非空凸集，且 $0\notin K$；0 可以是其边界点，则存在某个 $\lambda\in\mathbf{R}^n$，使得

$$\lambda'x\geqslant 0, \quad \text{对所有 } x\in K \text{ 成立,}$$

且至少存在一个 $x\in K$ 使得该不等式严格成立.

证明 如果 0 不是边界点，令 $m=\inf\limits_{x\in K}\|x\|>0$，则下确界 m 可在某个 $0\neq x_0\in\overline{K}$ 达到. 因为 $\overline{K}$ 是闭凸集，因此对任意的 $\alpha\in[0,0]$，有

$$\|x_0+\alpha(x-x_0)\|\geqslant\|x_0\|, \quad \text{对所有 } x\in K\subset\overline{K} \text{ 成立,}$$

设 $\lambda:=x_0$，令 $\alpha\to 0$，则有 $\lambda'x\geqslant\|\lambda\|^2>0$.

如果 0 是 K 的边界点，基本思想是将集合 K 整体平移一段距离，然后用上面的分离定理. 为了便于阅读，下面给出证明的细节，其过程有点繁琐.

首先证明 $\overline{K}$ 是 $\mathbf{R}^n$ 的真子集，即 $\overline{K}\neq\mathbf{R}^n$. 因此必须找到 $y\in\mathbf{R}^n\setminus\overline{K}$. 令 $\{x_1,\cdots,x_m\}$ 是 K 的一个极大线性无关向量组，即 $K=\text{span}\{x_1,\cdots,x_m\}$，$K$ 中任意向量都可以由 $x_1,\cdots,x_m$ 线性表出，但需要注意的是，不是 $x_1,\cdots,x_m$ 的所有线性组合都在 K 中. 例如，考虑如下向量：

$$y=-(x_1+\cdots+x_m).$$

假设 $y\in\overline{K}$，则存在点列 $\{y_k\}\subset K$ 使得 $y_k\to y$，且每一个 y_k 均可表示成 $x_1,\cdots,x_m$ 的线性组合，即存在实系数 λ_{ki}，使得

$$y_k=\sum_{i=1}^m\lambda_{ki}x_i.$$

因为 $y_k \to y$，则当 $k\to\infty$ 时，有 $\lambda_{ki}\to -1$ 成立. 所以必有足够大的 $k_0\in \mathbf{N}$，使得 $\lambda_{k_0,i}<0$，$i=1$，…，m. 从而得到

$$y_{k_0}+\sum_{i=1}^{m}(-\lambda_{k_0,i})x_i=0.$$

系数 1，$-\lambda_{k_0,1}$，…，$-\lambda_{k_0,m}$都是正数，它们的和不一定为 1，但只要将各项都分别除以它们的和 $1-\sum\limits_{i=1}^{m}\lambda_{k_0,i}$，则得到 y_{k_0}，x_1，…，$x_m\in K$ 的一个凸组合，且值为 0，这与 $0\notin K$ 矛盾，因此 $\overline{K}\neq \mathbf{R}^n$.

54

又因为 0 是 K 的一个边界点，则有 $\inf\limits_{x\in K}\|x\|=0$. 从而可以选取某个点列 $\{z_k\}\subset \mathbf{R}^n$，使得

$$z_k\to 0,k\to\infty,\quad \inf_{x\in K}\|x-z_k\|>0. \tag{2.5}$$

事实上，为了构造这样一个点列，设 x 是 K 的一个内点，记 $T(x)=-x$ 是原集关于原点的反射集，则 $y=T(x)\notin K$，否则连接 x 与 y 的线段属于 K，这与 $0\notin K$ 矛盾. 取 $\varepsilon>0$ 足够小，使得$B(x,\varepsilon)\subset K$，由算子 T 的连续性，集合

$$V=T(B(x,\varepsilon))=\{T(z):z\in B(x,\varepsilon)\}$$

是 $y=-x$ 的一个开邻域，并且可以选取 ε 使得 $V\cap K=\varnothing$. 因此，y 不是 K 的边界点. 则可以断言连接 y 和 0 线段上的所有的点 u 必满足 $\inf\limits_{x\in K}\|u-z\|>0$. 事实上，如果 u 是满足 $\inf\limits_{x\in K}\|u-z\|=0$ 的一个点，则可找到 u 的附近某个 $\tilde{u}\in K$，使得连接$\tilde{u}$和原点的直线可以分割 $B(x,\varepsilon)$球，但这意味着 $0\in K$，与已知矛盾. 找到点 y 后，可以构造序列 $z_k=y/k$，$k\in\mathbf{N}$，使之满足(2.5)式. 由构造知：集合 $K_k=K-z_k$ 均满足 $\inf\limits_{x\in K_k}\|x\|=\inf\limits_{x\in K}\|x-z_k\|>0$. 由定理 2.4.1 知：存在 $\lambda_k\in\mathbf{R}^n$ 和实数 $a_k>0$ 使得

$$\lambda_k' x\geqslant a_k,\quad 对所有\ x\in K_k\ 成立.$$

不失一般性，设$\|\lambda_k\|=1$，则点列$\{\lambda_k\}$属于紧集 $S^{n-1}=\{x\in\mathbf{R}^n:\|x\|=1\}$. 因此存在一个收敛子列$\{\lambda_{k_l}:l\in\mathbf{N}\}\subset\{\lambda_k:k\in\mathbf{N}\}$，其极限为 λ. 因为 $z_l\to 0$，由引理 2.4.2 知

$$\lambda' x=(\lim_{l\to\infty}\lambda_{k_l})'(\lim_{l\to\infty}(x-z_{k_l}))\geqslant 0$$

对所有 $x\in K$ 成立.

最后只需证明 λ 与 K 不正交即可(即 K 至少有一个元素使得定理中的不等式严格成立). 如若不然，则对任意的 $x\in K$，均有 $\lambda' x=0$，则可推出 $K\subset U=\{y\in\mathbf{R}^n:y\perp\lambda\}\subsetneq\mathbf{R}^n$，这与 K 不属于任意一个线性子空间矛盾. 因此至少可以找到一个 $x_0\in K$，使得 $\lambda' x_0>0$. ■

由线性代数知道，对任意的线性子空间 $U\subset\mathbf{R}^n$，必可以找到使得 $\lambda\in\mathbf{R}^n$，使得对任意的 $x\in U$，均有 $\lambda' x=0$ 成立. 下面的定理表明，只要 $U\cap K=\varnothing$，则可选择 λ 使得线性函数$\lambda(x)=\lambda' x$ 在紧集 K 上恒为正. 回顾一个记号

$$A-B=\{x\in\mathbf{R}^n:x=a-b,其中,a\in A,b\in B\},$$

下列引理是非常简单的.

引理 2.4.4　*以下命题成立.*

(ⅰ) 如果 A 和 B 均是凸集，则 $A-B$ 也是凸集.

(ⅱ) 如果 A 是紧集，B 是闭集，则 $A-B$ 是闭集.

证明 (ⅰ) 由定义直接得到.

(ⅱ) 设当 $n\to\infty$ 时，$x_n=a_n-b_n\to x$，其中 $\{a_n\}\subset A$，$\{b_n\}\subset B$，下面证明 $x\in A-B$，即 $x=a-b$，且 $a\in A$，$b\in B$.

55

因为 A 是紧集，即有界闭集，则必可取 $\{a_k\}$ 中的某个子列 $\{a_{n_k}: k\in\mathbf{N}\}$ 使得当 $k\to\infty$ 时，$a_{n_k}\to a$，从而当 $k\to\infty$ 时，$b_{n_k}=a_{n_k}-x_{n_k}\to a-x$. 又因为 B 是闭集，其极限值 $b:=a-x$ 必在集合 B 内. 综上所述，有 $x=a-b\in A-B$. ■

定理 2.4.5 设 $U\subset\mathbf{R}^n$ 是一个线性子空间，$\varnothing\neq K\subset\mathbf{R}^n$ 为凸紧集，且 $U\cap K=\varnothing$. 则存在 $\lambda\in\mathbf{R}^n$，使得：

$$\lambda' x = 0, \quad 对所有\ x\in U\ 成立;$$

且

$$\lambda' x > 0, \quad 对所有\ x\in K\ 成立.$$

证明 易知 $K-U$ 是闭凸集. 而由 $U\cap K=\varnothing$ 知 $0\notin K-U$. 由定理 2.4.1 知，存在 $\lambda\in\mathbf{R}^n$ 和实数 $a>0$，使得

$$\lambda' x \geqslant a, \quad 对所有\ x\in K-U\ 成立;$$

由此得到

$$(*) \quad \lambda'(y-x)=\lambda' y-\lambda' x\geqslant a, \quad 对所有\ y\in K\ 和\ x\in U\ 成立,$$

此时，假设存在某个 $x\in U$ 使得 $\lambda' x\neq 0$. 因为 $\lambda'(kx)=k\cdot\lambda' x$，对所有 $k\in\mathbf{Z}$ 成立，所以必能找到足够大的 k 使得 $\lambda' y-\lambda'(kx)<a$，这与 $(*)$ 式矛盾. 因此，对任意的 $x\in U$ 均有 $\lambda' x=0$；同时对所有的 $y\in K$ 有 $\lambda' y\geqslant a>0$. ■

2.5 无套利与鞅测度的关系

回顾一下前面讲过的一个概念：可测空间 $(\Omega, \mathcal{F})$ 上的两个概率测度 P 和 Q 称为等价的，记为 $P\sim Q$，如果它们有同样的零测集，即对任意的 $A\in\mathcal{F}$，均有 $P(A)=0\Leftrightarrow Q(A)=0$. 如 $P\sim Q$ 成立，则 Radon-Nikodym 导数 $\mathrm{d}P/\mathrm{d}Q$（P 关于 Q 的 R-N 导数）存在且是恒正的.

定义 2.5.1（等价鞅测度） 考虑一个金融市场，所在概率空间为 $(\Omega, \mathcal{F}, P)$，无风险利率为 r，风险资产为 $\{S_{t1}: t=1, 2\}$.

(ⅰ) 称 $\mathcal{F}$ 上的概率测度 P^* 为**鞅测度**，如果

$$S_{0i}=E^*(S_{1i}^*)=E^*\left(\frac{S_{1i}}{1+r}\right), \quad i=1,\cdots,d.$$

以后如不作特别说明，则 E^* 表示在概率测度 P^* 下求期望.

(ⅱ) 称 $(\Omega, \mathcal{F})$ 上的概率测度 P^* 为 P 的**等价鞅测度**，如果 P^* 是一个鞅测度，且有 $P^*\sim P$.

56

(ⅲ) 用

$$\mathcal{P}=(P^*: P^*\ 是\ P\ 的等价鞅测度)$$

表示 P 的所有等价鞅测度的集合.

注 2.5.2 后面将用到下列记号:

$$E^*(X) = E_{P^*}(X) = \int_\Omega X \mathrm{d}P^*.$$

注 2.5.3

(i) 由 $S_{0i}^* = S_{0i}$ 得前面定义中的条件等价于

$$S_{0i}^* = E^*(S_{1i}^*), \quad i = 1, \cdots, d.$$

即 d 维贴现价格过程 $\{S_{ti}^*: t=0, 1\}$ 在测度 P^* 下是鞅. 鞅在数理金融中起着非常重要的作用, 下一小节将详细讨论它.

(ii) 由等价鞅测度的定义知: 集合 $\mathcal{P}$ 与真实概率测度 P 有关.

(iii) P^* 是一个鞅测度的充分必要条件是购买的第 i 种资产在 P^* 下的期望贴现收益为 0, 即:

$$E^*(G_{1i}^*) = E^*(S_{1i}^* - S_{0i}) = 0, \quad i = 1, \cdots, d.$$

鞅测度在数理金融中有着非常重要的作用. 在第 1 章中, 用期望表示随机支付 X 的价格时, 它忽略了与 X 相关的风险, 因此把鞅测度 P^* 与真实概率测度 P 区别开是非常重要的. 一般来说, 期望 $E_P(S_{1i}/(1+r))$ 不是实际市场价格, 但是在风险将被忽略的风险中性鞅测度 P^* 下, 随机贴现支付是公平价格. 其中 P^* 是在风险中性的假设下, 对市场可能出现的每一种情况 $\omega \in \Omega$, 在每个事件 $A \subset \Omega$ 上取值.

下面引入期权定价的主要结论: 市场无套利等价于存在一个等价的鞅测度 P^*, 即 $\mathcal{P} \neq \varnothing$. 由于在任意的概率空间中证明这个结论是非常复杂的, 所以先考虑有限概率空间的特殊情况. 为此, 作如下假定:

$$\begin{aligned} &\Omega = \{\omega_1, \cdots, \omega_n\}, \quad \text{对某个 } n \in \mathbf{N}, \\ &\mathcal{F} = \mathrm{Pot}(\Omega), \\ &P(\{\omega\}) > 0 \quad \forall\, \omega \in \Omega. \end{aligned}$$

57

任一随机变量 $Z: \Omega \to \mathbf{R}$ 可以映射为一个向量 $(Z(\omega_1), \cdots, Z(\omega_n))'$. 反之亦然, 映射

$$\mathbf{R}^n \ni (y_1, \cdots, y_n)' \mapsto Z, Z(\omega_i) = y_i, \quad i = 1, \cdots, n,$$

也定义了一个随机变量 Z. 显然, 该映射是一一映射. 所以, 可以将一个随机变量表示成向量的形式, 可以用一个矩阵来表示随机向量. $((d+1) \times n)$ 维支付矩阵为

$$\boldsymbol{D} = \begin{bmatrix} S_{10}(\omega_1), & \cdots, & S_{10}(\omega_n) \\ \vdots & & \vdots \\ S_{1d}(\omega_1), & \cdots, & S_{1d}(\omega_n) \end{bmatrix}.$$

其中, 第 i 行表示第 i 项投资, 第 j 列表示在某种市场行情下整个组合的支付. 则有如下关系式

$$\varphi' S_1 = \boldsymbol{D}' \varphi,$$

如上面所解释的那样, 向量 $(\varphi' S_1(\omega_1), \cdots, \varphi' S_1(\omega_n))'$ 表示随机变量 $\varphi' S_1$.

定理 2.5.4 (Ω 有限条件下资产定价的基本定理) 设 $\Omega = \{\omega_1, \cdots, \omega_n\}$, 如果金融市场是无套利的, 则存在一个向量 $\psi \in \mathbf{R}^n$, 满足: 它的每个分量 ψ_j 都是正的且 $S_0 = \boldsymbol{D}\psi$, 使

得按下列方式

$$P^*(\{\omega_i\})=\frac{\psi_i}{\sum_{j=1}^{n}\psi_j},\quad i=1,\cdots,n,$$

确定的$\mathcal{F}$上的测度P^*为P的等价鞅测度．即$E^*(S_1^*)=S_0$成立，且$\sum_{j=1}^{n}\psi_j=(1+r)^{-1}$为贴现因子．

证明 考虑集合

$$U=\{(-\varphi' S_0,(\varphi' S_1)')':\varphi\in\mathbf{R}^{d+1}\}\subset\mathbf{R}\times\mathbf{R}^n=\mathbf{R}^{n+1},$$

显然，U是$\mathbf{R}^{n+1}$的线性子空间．令

$$M=\{(y_0,\cdots,y_n)'\in\mathbf{R}^{n+1}:y_i\geqslant 0,i=0,\cdots,n;y_j>0,\text{对某个 }j=1,\cdots,n\}.$$

则有$0\notin M$．易知：

$$\text{市场无套利}\Leftrightarrow U\cap M=\varnothing.$$

设

$$K=\left\{(y_0,\cdots,y_n)'\in M:\sum_{i=0}^{n}y_i=1\right\}.$$

K是M中的一个非空、凸的、紧子集，且$U\cap K=\varnothing$．由定理2.4.5，可以对K和U进行分离，即

$$\exists\lambda=(\lambda_0,\lambda_1,\cdots,\lambda_n)\in\mathbf{R}^{n+1},$$

使得

$$\lambda' x=0,\quad\forall x\in U,$$
$$\lambda' x>0,\quad\forall x\in K.$$

见图2-1给出的描述．

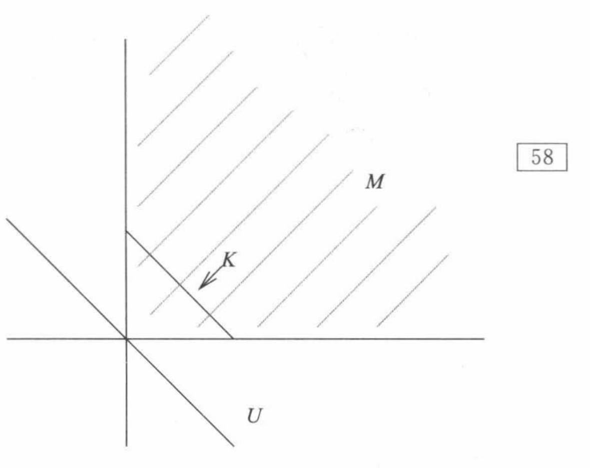


图2-1 图中U是投资组合φ对应的投资收益$\varphi' S_1$的线性子空间，如果$\varphi' S_0\leqslant 0$，M对应一个套利机会，K是M的凸紧子集，根据定理2.4.5，K和U是可分离的

因为$e_j=(0,\cdots,0,1,0,\cdots,0)'\in K$，所以有$\lambda_j=\lambda' e_j>0$，$j=0,\cdots,n$．设$\widetilde{\lambda}=(\lambda_1,\cdots,\lambda_n)'$，由$U$的定义知

$$\begin{aligned}&-\lambda_0\varphi' S_0+\widetilde{\lambda}'\underbrace{\varphi' S_1}_{=\mathbf{D}'\varphi}=0 && \forall\varphi\in\mathbf{R}^{d+1}\\ \Leftrightarrow\quad&-\lambda_0 S_0'\varphi+(\boldsymbol{D}\widetilde{\lambda})'\varphi=0 && \forall\varphi\in\mathbf{R}^{d+1}\\ \Leftrightarrow\quad&(\boldsymbol{D}\widetilde{\lambda})'\varphi=(\lambda_0 S_0)'\varphi && \forall\varphi\in\mathbf{R}^{d+1}.\end{aligned}$$

所以得到

$$\boldsymbol{D}\widetilde{\lambda}=\lambda_0 S_0\quad\Leftrightarrow\quad S_0=\boldsymbol{D}\left(\frac{\widetilde{\lambda}}{\lambda_0}\right).$$

定义

$$\psi:=(\psi_1,\cdots,\psi_n)':=\frac{\widetilde{\lambda}}{\lambda_0}=\left(\frac{\lambda_1}{\lambda_0},\cdots,\frac{\lambda_n}{\lambda_0}\right)'\in\mathbf{R}^n.$$

58
59

则

$$p_j^* := P^*(\{\omega_j\}) := \frac{\psi_j}{\sum_{k=1}^{n}\psi_k}, \quad j=1,\cdots,n,$$

定义了$(\Omega, \mathcal{F})$上的一个与 P 等价的概率测度 P^*．事实上，对所有的 j 均有 $\lambda_j>0$ 包含了对所有的 j 均有 $\psi_j>0$．

现取 $\overline{\varphi}\in \mathbf{R}^{d+1}$ 满足

$$\boldsymbol{D}'\widetilde{\varphi} = \mathbf{1}\cdot(1+r),$$

则 $\overline{\varphi}$ 复制了银行账户的资产．因此 $\overline{\varphi}$ 的初始价格为 1，否则就会出现套利机会．所以

$$1 = \overline{\varphi}'S_0 = \overline{\varphi}'(\boldsymbol{D}\psi) = \psi'\boldsymbol{D}'\overline{\varphi} = \psi'\mathbf{1}\cdot(1+r) = (1+r)\sum_{j=1}^{n}\psi_j.$$

由此得到，$\sum_{j=1}^{n}\psi_j = (1+r)^{-1}$ 是模型中的贴现因子．最后，再次由 $S_0=\boldsymbol{D}\psi$ 知，对 $i=1, \cdots, d$，有

$$\begin{aligned} E^*(S_{1i}^*) &= E^*\left(\frac{S_{1i}}{1+r}\right) \\ &= \frac{1}{1+r}\sum_{j=1}^{n}S_{1i}(\omega_j)p_j^* \\ &= \frac{1}{1+r}\sum_{j=1}^{n}D_{ij}\psi_j\Big/\sum_{k=1}^{n}\psi_k \\ &= \sum_{j=1}^{n}D_{ij}\psi_j = (\boldsymbol{D}\psi)_i = S_{0i}. \end{aligned}$$

从而得到 $E^*(S_1^*)=S_0$． ■

接下来要将上一结论推广到一般的概率空间中去．作为准备，先复习一下下列简单事实．

引理 2.5.5 如果 $E(X\mathbf{1}_{\{X<0\}})\geqslant 0$，那么，$X\geqslant 0$(a. s.)．

证明 因为对任意 $\omega\in\Omega$ 均有 $X\mathbf{1}_{\{X<0\}})\leqslant 0$，由测度积分的单调性知，$\int X\mathbf{1}_{\{X<0\}})\mathrm{d}P\leqslant 0$，因此，$\int X\mathbf{1}_{\{X<0\}})\mathrm{d}P=0$，这等价于 $X\mathbf{1}_{\{X<0\}}=0$，P-a. s.，所以 $N=\{X\mathbf{1}_{\{X<0\}}\neq 0\}$，再由 $A=\{X<0\}\subset N$ 知 A 为 P-零测集，所以有 $P(X\geqslant 0)=1$． ■

定理 2.5.6(一般空间 Ω 中的资产定价基本定理)

60

(i) 若$\mathcal{P}\neq\varnothing$，则金融市场是无套利的．

(ii) 如果市场无套利，则存在等价鞅测度 $P^*\in\mathcal{P}$，使得 $\mathrm{d}P^*/\mathrm{d}P$ 是有界的．

证明 (i) 用反证法．设存在测度 $P^*\in\mathcal{P}$，且市场上存在一个套利机会的资产组合 $\varphi\in\mathbf{R}^{d+1}$，满足 $\varphi'S_1\geqslant 0$ 且 $E(\varphi'S_1)>0$．由 P 和 P^* 的等价性知 $E^*(\varphi'S_1)>0$，所以

$$\sum_{i=0}^{d}\varphi_i E^*\left(\frac{S_{1i}}{1+r}\right) = E^*\left(\frac{\varphi'S_1}{1+r}\right) > 0.$$

又因为 P^* 是鞅测度，$S_0=E^*(S_1/(1+r))$，所以有

$$\varphi' S_0 = \sum_{i=0}^{d} \varphi_i S_{0i} = \sum_{i=0}^{d} \varphi_i E^* \left(\frac{S_{1i}}{1+r}\right) > 0.$$

这表明 φ 不是套利组合，与假设矛盾.

（ⅱ）由 d 个风险资产对应的 d 维贴现收益过程的定义知

$$G_1^* = (G_{11}^*, \cdots, G_{1d}^*), \quad G_{1i}^* = S_{1i}^* - S_{0i}, \quad i = 1, \cdots, d.$$

根据引理 2.3.4，无套利条件等价于

$$(*) \quad \forall \varphi \in \mathbf{R}^d: \sum_{i=1}^{d} \varphi_i G_{1i}^* \geqslant 0, P\text{-a. s.} \Rightarrow \sum_{i=1}^{d} \varphi_i G_{1i}^* = 0, P\text{-a. s.}$$

下面证明上述结论能推出等价鞅测度 $P^* \sim P$ 的存在性. 首先注意到 $P^* \sim P$ 是鞅测度的充分必要条件是

$$E^*(S_1^*) = S_0 \quad \Leftrightarrow \quad E^*(G_1) = 0.$$

为了简化记号，以下记 $G=G_1^*$ 为 1 时刻的贴现收益价格，该记号并不与标准记号冲突，因为它只在本证明中使用. 首先考虑 G 是 P 可积的情况，即 $E|G|<\infty$. 定义

$$\mathcal{K} = \{Q: Q \text{ 为概率测度且 } Q \sim P, \mathrm{d}Q/\mathrm{d}P \text{ 有界}\}.$$

因为 $P\in\mathcal{K}$，所以$\mathcal{K}\neq\varnothing$. 对所有的 $Q\in\mathcal{K}$，如果 $\mathrm{d}Q/\mathrm{d}P\leqslant C<\infty$，则

$$E_Q|G| = \int |G(\omega)| \frac{\mathrm{d}Q}{\mathrm{d}P}(\omega)\mathrm{d}P(\omega) \leqslant CE_P|G| < \infty,$$

61

因此，G 在所有测度 Q 下也是可积的. 可以直接验证$\mathcal{K}$是一个凸集. 考虑 G 在所有 $Q\in\mathcal{K}$下的期望构成的集合：

$$\mathcal{E} = \{E_Q(G): Q \in \mathcal{K}\} \subset \mathbf{R}^d.$$

如果 $0\in\mathcal{E}$，那么等价鞅测度 P^* 存在. 事实上，此时满足 $E^*(G)=0\in\mathcal{E}$的任何 $P^*\in\mathcal{K}$都是等价鞅测度. 因此，只需证明 $0\in\mathcal{E}$. 如若不然，则 $0\notin\mathcal{E}$.

首先注意到$\mathcal{E}$是一个凸集. 事实上，设 x_1，$x_2\in\mathcal{E}$，$\alpha\in[0, 1]$，那么存在 Q_1，$Q_2\in\mathcal{K}$使得 $x_i=E_{Q_i}(G)$，$i=1, 2$. 由 $\mu=\alpha Q_1+(1-\alpha)Q_2\in\mathcal{K}$知

$$\alpha x_1 + (1-\alpha)x_2 = \alpha\int G\mathrm{d}Q_1 + (1-\alpha)\int G\mathrm{d}Q_2$$

$$= \int G\mathrm{d}[\alpha Q_1 + (1-\alpha)Q_2] = \int G\mathrm{d}\mu \in \mathcal{E},$$

进一步，由 $P\in\mathcal{K}$得$\mathcal{E}\neq\varnothing$.

由定理 2.4.3 知，可以找到一个分离$\mathcal{E}$和$\{0\}$的向量 $\varphi\in\mathbf{R}^d$. 即，

$$\varphi' x \geqslant 0 \quad \text{对所有 } x \in \mathcal{E} \text{成立.}$$

且

$$\varphi' x_0 > 0 \quad \text{对某个 } x_0 \in \mathcal{E} \text{成立.}$$

因为对任意的 $Q\in\mathcal{K}$，均有 $x=E_Q(G)\in\mathcal{E}$，所以

$$E_Q(\varphi' G) \geqslant 0, \quad \text{对所有 } Q \in \mathcal{K} \text{成立} \tag{2.6}$$

且

$$E_{Q_0}(\varphi' G) > 0 \quad \text{对某个 } Q_0 \in \mathcal{K} \text{成立}.$$

又因为 $Q_0 \sim P$，$E_{Q_0}(\varphi' G) > 0$，因此得到 $P(\varphi' G > 0) > 0$，如能证明也有 $\varphi' G \geqslant 0$，P-a. s.，则由(*)式得 $\varphi' G = 0$(P-a. s.)，这与 $P(\varphi' G > 0) > 0$ 矛盾，那么结论得证.

下面证明 $\varphi' G \geqslant 0$(P-a. s.)，为此只需证明 $E_P(\varphi' G \mathbf{1}_{\{\varphi' G < 0\}}) \geqslant 0$ 即可，这将用到引理 2.5.5. 事实上，可以把 $\mathbf{1}_{\{\varphi' G < 0\}}$ 看作新测度 Q 的有界 P-密度，则有 $Q \in \mathcal{K}$，由(2.6)式得到 $E_Q(\varphi' G) \geqslant 0$，然而，要使 Q 为概率测度，需要除以 $P(\varphi' G < 0)$，这是不可能的.

62

接下来，用一列函数去逼近示性函数 $\mathbf{1}_{\{\varphi' G < 0\}}$，这些函数的支集与$\{\varphi' G \geqslant 0\}$有非空交集. 为此，定义

$$f_n = (1 - 1/n)\mathbf{1}_{\{\varphi' G < 0\}} + (1/n)\mathbf{1}_{\{\varphi' G \geqslant 0\}}, \quad n \in \mathbf{N}.$$

函数列$\{f_n\}$有下列性质，其证明留给读者作为练习.

(i) $0 < f_n \leqslant 1$，$n \geqslant 2$.

(ii) $f_n \to \mathbf{1}_{\{\varphi' G < 0\}}$，$n \to \infty$，$\omega$-逐点的.

(iii) $\int f_n \mathrm{d}P = (1 - 1/n)P(\varphi' G < 0) + (1/n)P(\varphi' G \geqslant 0) > 0$，对所有 $n \in \mathbf{N}$.

(iv) $\omega \mapsto \widetilde{f}_n(\omega) = f_n(\omega) / \int f_n \mathrm{d}P$，$\omega \in \Omega$，定义有界的 P-密度.

令 $Q_n = \widetilde{f}_n \mathrm{d}P$ 为对应的概率测度，即

$$Q_n(A) = \int_A \widetilde{f}_n \mathrm{d}P, \quad A \in \mathcal{F}.$$

由性质(1)知 $Q_n \in \mathcal{K}$，$n \geqslant 2$. 根据(2.6)式有

$$0 \leqslant E_{Q_n}(\varphi' G) = E_P(\varphi' G f_n) / \int f_n \mathrm{d}P.$$

对上式右边的分子应用控制收敛定理，得

$$E_P(\varphi' G \mathbf{1}_{\{\varphi' G < 0\}}) = \lim_{n \to \infty} E_P(\varphi' G f_n) \geqslant 0.$$

于是由引理 2.5.5 得到 $\varphi' G \geqslant 0$(P-a. s.).

最后，考虑 $E_P|G| = \infty$的情形，因为

$$0 < c = E\left(\frac{|G|}{|G| + 1}\right) \leqslant 1.$$

令 P'为 P-密度是

$$f(\omega) = \frac{\widetilde{c}}{1 + |G(\omega)|}$$

的概率测度，其中$\widetilde{c}$的选取确保 $P'(\Omega) = \int f(\omega)\mathrm{d}P(\omega) = 1$. 因此 $P' \sim P$，且

63

$$E_{P'}|G| = E_P\left(\frac{|G|}{|G| + 1}\widetilde{c}\right) < \infty.$$

所以 G 在 P'下是可积的. 并可以断定存在鞅测度 P^* 满足 $P^* \sim P'$，且有有界密度. 显然，也有 $P^* \sim P$，而且 P^* 的 P-密度也是有界的. 事实上，

$$\frac{\mathrm{d}P^*}{\mathrm{d}P} = \frac{\mathrm{d}P^*}{\mathrm{d}P'} \cdot \frac{\mathrm{d}P'}{\mathrm{d}P} \leqslant C,$$

其中 C 为某个固定的常数. ■

下面给出与结论相关的两个例子.

▶**例 2.5.7**(单资产的二叉树模型) 回顾前面所介绍的简单二叉树模型，其样本空间可取为 $\Omega=\{+, -\}$，实际概率测度 P 由 $p_+=P(\{+\})\in(0, 1)$ 确定. 资产价格或以概率 p_+ 从 S_{01} 上涨到 $S_{11}=uS_{01}$，或以概率 $p_-=1-p_+$ 下降到 dS_{01}. 市场无套利当且仅当 $\mathcal{P}\neq\varnothing$. 显然，$P^*\in\mathcal{P}$ 由 $p^*=P^*(\{+\})$ 唯一确定，且 $P^*\sim P$ 等价于 $p^*\in(0, 1)$. 在 1.5.2 节中已经讨论了下列鞅方程

$$E^*\left(\frac{S_{11}}{1+r}\right)=S_{01},$$

得到其解为

$$p^*=\frac{1+r-d}{u-d},\quad 1-p^*=\frac{u-(1+r)}{u-d}.$$

所以

$$0<p^*<1 \quad\Leftrightarrow\quad 0<1+r-d<u-d \quad\Leftrightarrow\quad d<1+r<u.$$

最后，计算 P^* 关于 P 的 Radon-Nikodym 导数，

$$\frac{\mathrm{d}P^*}{\mathrm{d}P}(\{+\})=\frac{P^*(\{+\})}{P(\{+\})}=\frac{p^*}{p}=\frac{1+r-d}{u-d}\cdot\frac{1}{p},$$

$$\frac{\mathrm{d}P^*}{\mathrm{d}P}(\{-\})=\frac{1-p^*}{1-p}=\frac{u-(1+r)}{u-d}\cdot\frac{1}{1-p},$$

这就确定了映射 $\frac{\mathrm{d}P^*}{\mathrm{d}P}:\Omega\to\mathbf{R}$. ◀

▶**例 2.5.8**(离散金融市场) 设 $\Omega=\{\omega_1, \cdots, \omega_n\}$ 表示市场上所有可能出现的 n 种情况，不妨设 $P(\{\omega_i\})>0$，对 $i=1, \cdots, n$ 均成立(否则可以将其从空间 Ω 中剔除). 每一个鞅测度均由概率向量 $p^*=(p_1^*, \cdots, p_n^*)'\in\mathbf{R}^n$ 确定，其中 $p_1^*, \cdots, p_n^*\geqslant 0$ 且 $\sum_i p_i^*=1$. P^* 是等价鞅测度当且仅当对 $i=1, \cdots, n$，均有 $p_i^*\in(0, 1)$. 需要解方程组：

$$S_{0i}=E^*\left(\frac{S_{1i}}{1+r}\right)=\sum_{j=1}^{n}p_j^*\frac{S_{1i}(\omega_j)}{1+r},\quad i=1,\cdots,d,$$

64

引入支付矩阵 $\boldsymbol{D}=(S_{1i}(\omega_j))_{i=0,\cdots,d,j=1,\cdots,n}$，则方程组等价于

$$(1+r)S_0=\boldsymbol{D}p^*.$$

因此，p^* 是下列线性约束方程组的解：

$$\boldsymbol{D}x=(1+r)S_0,\quad x>0,\quad \mathbf{1}'x=1,$$

该方程组由 d 个方程和 n 个未知量构成. ◀

2.6 未定权益的无套利定价

本节介绍衍生产品无套利定价的基本理论，或更为广泛的未定权益的定价. 特别地，本节将证明如果存在一个等价鞅测度，则未定权益的对冲风险的资产组合的价格是唯一的.

已经知道，任意一个资产组合 $\varphi\in\mathbf{R}^{n+1}$ 在 $t=1$ 时刻会产生随机支付 $V(\varphi)=\varphi' S_1$. 考虑集合

$$\mathcal{V}=\{\varphi' S_1:\varphi\in\mathbf{R}^{d+1}\},$$

它是随机资产组合的支付集. 易知，$\mathcal{V}$是所有随机变量 $\Omega\to\mathbf{R}$ 的集合(以后记此集合为 L)的一个子集. 下列结论是显然的.

引理 2.6.1 $\mathcal{V}=\mathrm{span}(S_{10},\cdots,S_{1d})$是一个线性空间，且有 $1\in\mathcal{V}$.

由定义知，对任意的支付 $V\in\mathcal{V}$，都存在资产组合 $\varphi\in\mathbf{R}^{n+1}$，使其对应的支付为 V，即 $V=V(\varphi)$.

定义 2.6.2 设 $X:\Omega\to\mathbf{R}$ 为一个随机变量，如果存在某个 $\varphi\in\mathbf{R}^{n+1}$ 满足 $V(\varphi)=X$，即 $X\in\mathcal{V}$，则称 φ 为一个**对冲**.

确定支付 $X\in\mathcal{V}$的合理价格使其为对冲的初始资本是非常必要的.

定义 2.6.3 无套利市场中，定义随机支付 $X\in\mathcal{V}$的价格为

$$\pi(X)=\varphi' S_0,$$

其中 $\varphi\in\mathbf{R}^{n+1}$ 为一个对冲组合，即 $V(\varphi)=X$. 对应的线性映射 $\pi:\mathcal{V}\to\mathbf{R}$ 称为**线性定价法则**.

因为对冲可能是不唯一的，所以 $\pi(X)$是否为正确的定价还须检验. 但是，如果存在等价鞅测度，则该定义是确定的.

定理 2.6.4 如果市场存在等价鞅测度，那么，由任意两个资产组合 φ_1，$\varphi_2\in\mathbf{R}$ 得到的随机支付 $X\in\mathcal{V}$的价格是一样的，即 $\varphi_1' S_0=\varphi_2' S_0$.

65

证明 显然，$\delta=\varphi_1-\varphi_2$ 满足 $V(\delta)=(\varphi_1-\varphi_2)' S_1=0$. 因此，对任意的等价鞅测度 $P^*\in\mathcal{P}$，由 $P(V(\delta)=0)=1$ 知 $P^*(V(\delta)=0)=1$. 于是

$$0=E^*(V(\delta))=E^*((\varphi_1-\varphi_2)' S_1),$$

进一步有

$$0=E^*((\varphi_1-\varphi_2)' S_1^*)=(\varphi_1-\varphi_2)' E^*(S_1^*)=(\varphi_1-\varphi_2)' S_0. \quad \blacksquare$$

下面给出衍生产品和未定权益的严格数学定义. 前面所讲的金融工具均为它们的价值过程，经典组合 $V_t=\varphi' S_t$ 是资产的一个线性函数.

定义 2.6.5 设随机变量 $X:\Omega\to\mathbf{R}$ 满足 $0\leqslant X<\infty$(P-a. s.)，且关于 $\sigma(S_{11},\cdots,S_{1d})$ 可测，即存在某个 Borel 可测函数 f，使得 X 可以表示为

$$X=f(S_{11},\cdots,S_{1d}),$$

则称 X 为一个**衍生产品**或**衍生资产**.

注 2.6.6 对任意一项资产 X，可以取使 $X=f(S_{1i_1},\cdots,S_{1i_r})$在下列意义下最小的 f，即 $i_1,i_2,\cdots,i_r\in\{1,\cdots,d\}$且 r 的值最小. 称 $S_{1i_1},\cdots,S_{1i_r}$ 为**原生资产**或**标的资产**.

需要指出的是，以上给出的衍生产品的定义并没有覆盖金融市场上的所有期权和权证. 注意到衍生品 X 关于$\mathcal{F}_1=\sigma(S_{11},\cdots,S_{1d})$可测，但一般来说$\mathcal{F}_1\subsetneq\mathcal{F}$，对$\mathcal{F}_1\neq\mathcal{F}$的情形，有下的定义.

定义 2.6.7 任意一个(非负的)随机变量 $X:(\Omega,\mathcal{F})\to(\mathbf{R},\mathcal{B})$称为一个**未定权益**.

▶**例 2.6.8**(气候衍生品，能源期权) 因为能源消费和平均气温有关，因而交易市场

上出现了天气或气候衍生品．该衍生品被能源公司广泛用来对冲或减少风险．不论看涨或看跌，这些衍生品合约均在场外进行交易．由于大家一般都在18℃的分界点开暖气或空调，因此一般以18℃作为衍生品合约的平均温度阈值．平均气温大于18℃的天数称为所谓的采暖度天数(HDD)，而气温低于18℃的天数称为制冷度天数(CDD)，天气衍生品通常用HDD和CDD数来定义． ◀

按L_p空间术语，未定权益是P-几乎必然有界的随机变量组成的空间

$$L_0 = L_0(\Omega,\mathcal{F},P) = \{X:(\Omega,\mathcal{F},P) \to (\mathbf{R},\mathcal{B}):|X| < \infty, P\text{-a. s.}\}$$

66

中的元素．下面还将考虑下列空间：

$$L_\infty = L_\infty(\Omega,\mathcal{F},P) = \{X:(\Omega,\mathcal{F},P) \to (\mathbf{R},\mathcal{B});\|X\|_\infty < \infty\}$$

它是由一致有界的随机变量组成的空间，其中，范数$\|X\|_\infty$定义如下：

$$\|X\|_\infty = \inf\{c > 0:P(|X| > c) = 0\},$$

即$\|X\|_\infty$是几乎必然不被$|X|$超过的所有数的最小值．显然$L_\infty \subset L_0$，但下例表明，$L_0 \not\subset L_\infty$．

例 2.6.9 令$\Omega=[0,1]$，$P=\lambda$为勒贝格测度，$\mathcal{F}=\mathcal{B}_{|[0,1]}$表示$[0,1]$上的Borel-$\sigma$代数．定义$X(\omega)=1/\omega$，$\omega\in\Omega$，规定$1/0=\infty$．因为$\{X=\infty\}=\{0\}$是$\lambda$-零测集，且$P(|X|<\infty)=1$，所以$X\in L_0([0,1],\mathcal{B}_{|[0,1]},\lambda)$．显然，对任意的$c>0$，集合$\{|X|>c\}=[0,1/c]$的勒贝格测度为正，所以$\|X\|_\infty=\inf\{c>0:P(|X|>c)\}=\infty$，得到$L_0\not\subset L_\infty$．

判断一个未定权益定价是否合理的基本思想是：把未定权益作为一项新的资产进行交易时，要求该价格不会带来套利机会．

定义 2.6.10 称$\pi(C)$为未定权益C的公平价格，如果由下列$d+2$种资产构成的市场是无套利的：

$$\{S_{ti}:t = 0,1\},\quad i = 0,\cdots,d+1,$$

其中

$$S_{0,d+1} = \pi(C),\quad S_{1,d+1} = C$$

是无套利的．习惯上用$\Pi(C)$表示所有无套利价格的集合．

设C为一个未定权益，$\pi(C)$为C的无套利价格．如果C的市场交易价是$\pi(C)$，则市场是无套利的．根据定义，存在等价鞅测度P^*使得

$$E^*(S_{1i}^*) = S_{0i},\quad i = 0,\cdots,d+1.$$

特别地，有如下定价公式：

$$\pi(C) = E^*(C^*).$$

上式表明，在无套利市场中，只要求得了等价鞅测度，则可以由风险中性定价原理求得未定权益的公平价格．

简单地说，就是：扩展市场的等价鞅测度进行的风险中性定价可以得到未定权益的所有无套利价格．

命题 2.6.11 设$\mathcal{P}\neq\varnothing$，则$\Pi(C)=\{E^*(C^*):P^*\in\mathcal{P},\ E^*(C^*)<\infty\}$．因此，套利的一个下界为$\pi_-(C)=\inf\{E^*(C^*):P^*\in\mathcal{P}\}$．如果对所有$P^*\in\mathcal{P}$，均有$E^*(C^*)<\infty$，则一个上界为$\pi_+(C)=\sup\{E^*(C^*):P^*\in\mathcal{P}\}$．

67

证明 设$\pi(C)$为扩展市场的无套利价格，则扩展市场存在一个等价鞅测度，使得

$$\pi(C)=E^*(C^*)\text{ 且 }S_{0i}=E^*(S_{1i}^*),\quad i=0,\cdots,d.$$

因此，$\Pi(C)\subset\{E^*(C^*):P^*\in\mathcal{P}$且$E^*(C)<\infty\}$. 反过来，如果存在某个$P^*\in\mathcal{P}$使得$x=E^*(C^*)$，显然有$S_{0i}=E^*(S_{1i}^*)$，$i=0$，…，$d$. 令$S_{0,d+1}=x$，$S_{1,d+1}=C$，则有$S_{0i}=E^*(S_{1i}^*)$，$i=0$，…，$d+1$. 所以扩展市场仍然是无套利的，$P^*\in\mathcal{P}$就是它的一个等价鞅测度. C可以由组合$\widetilde{\varphi}=(0,\cdots,0,1)'$复制，根据定理2.6.4，它的公平价格不依赖于$\widetilde{\varphi}$，而由$\widetilde{\varphi}'S_0=S_{0,d+1}=x$决定.

显然，C在扩展市场是可复制的. 如果C可以作为原始市场上的资产，C就可以看成是一个衍生品. 下面集中讨论这一点. ■

定义 2.6.12 未定权益C称为**可实现的**或**可复制的**，如果市场存在某个对冲组合$\varphi=(\varphi_0,\cdots,\varphi_d)'\in\mathbf{R}^{d+1}$使得

$$C=\varphi'S_1,$$

即$C\in\mathcal{V}$. 在这种情况下，C是一个衍生品.

不禁要问：无套利价格是唯一的吗？对可实现的未定权益，因为存在组合，所以答案是肯定的. 下面的命题概括了上述结论.

命题 2.6.13 设金融市场是无套利的，如果C是可复制的未定权益，则无套利价格是唯一的，且有$\pi_C=\varphi'S_0$成立，其中φ为任一复制资产组合.

现在考虑更有趣也更复杂的情形，未定权益是不可复制的. 此时无套利定价不再适用，它将用到泛函分析中的某些结论，这已经超出了本书的范围. 第一个结论是将定理2.4.1推广到赋范空间.

定理 2.6.14 设$Y\subset X$是赋范线性空间X的一个闭线性子空间，取其中某个$x_0\in X\setminus Y$，则存在满足$\|L\|=1$的线性泛函$L:X\to\mathbf{R}$，使得

$$L(y)=0\text{ 对所有的 }y\in Y,$$

且$L(x_0)=\mathrm{dist}(x_0,Y)$.

由定理2.6.14可以推出下列结论：首先，在线性子空间$Y_0=Y\oplus\mathrm{span}\{x_0\}$上定义线性泛函$L$：

68

$$L(y+tx_0)=t\cdot\mathrm{dist}(x_0,Y),\quad y\in Y,\quad t\in\mathbf{R}.$$

显然，对所有的$y\in Y$，有$L(y)=0$，且$L(x_0)=\mathrm{dist}(x_0,Y)$. 由Hahn-Banach定理，该线性泛函可以延拓到空间X上去.

令μ为一个σ有限测度，则$L_p=L_p(\mu)$，$1\leqslant p\leqslant\infty$空间是巴拿赫空间，它是一个完备赋范线性空间，即所有的柯西列均收敛. 由Hölder不等式知，对任意$X\in L_p$，$Y\in L_q$，有

$$\int XY\mathrm{d}\mu\leqslant\|X\|_p\|Y\|_q<\infty.$$

其中，$p\in[1,\infty)$，$q\in(1,\infty)$且$\frac{1}{p}+\frac{1}{q}=1$，并规定$q=\infty$时$p=1$. 因此，如果固定$Y\in L_q$，考虑映射

$$L_Y:L_p\to(\mathbf{R},\mathcal{B}),$$

其中

$$L_Y(X)=\int XY\mathrm{d}\mu,\quad X\in L_p,$$

则得到了一个定义在 L_p 上的线性连续泛函 L_Y. 又因为该泛函对任意的 $Y\in L_q$ 均有定义，所以映射

$$J:Y\mapsto L_Y,\quad 对 Y\in L_q$$

将 L_q 映射到它的对偶空间. 所以 J 是一个线性等距同构算子. 因此，可以认为 L_p 的对偶空间就是 L_q. 于是，所有线性连续泛函均有

$$L_Y(X)=\int XY\mathrm{d}\mu,\quad 对某个 Y\in L_q$$

的形式.

有了以上这些工具，现可以证明下列定理了.

定理 2.6.15 设未定权益 C 不可复制，且 $E^*(C)<\infty$，则无套利价格集是一个开区间.

证明 因为等价鞅测度集$\mathcal{P}$是一个凸集，因此 $\Pi(C)=\{E^*(C^*):P^*\in\mathcal{P}\}$也是凸集. 事实上，任取 P_1^*，$P_2^*\in\mathcal{P}$，$\varphi\in[0,1]$，定义测度 $\mu=\alpha P_1^*+(1-\alpha)P_2^*\in\mathcal{P}$，则有

$$\alpha\int C^*\mathrm{d}P_1^*+(1-\alpha)\int C^*\mathrm{d}P_2^*=\int C^*\mathrm{d}\mu\in\Pi(C).$$

因此，$\Pi(C)$是一个区间或空集. 下面证明 $\Pi(C)$是开区间，即对任意的 $E^*(C^*)$，存在 $\pi_-=\pi_-(C)$，$\pi_+=\pi_+(C)$，π_-，$\pi_+\in\Pi(C)$，且

$$\pi_-<E^*(C^*)<\pi_+.$$

因为 $E^*(S_{1i})=S_{0i}\in[0,\infty)$，所以，对所有的 $i=0,\cdots,d$，有 $S_{1i}\in L_1(P^*)$，那么，$\mathcal{V}=\mathrm{span}\{S_{10},\cdots,S_{1d}\}$是 $L_1(P^*)$的线性子空间. 因为 $C\in L_1(P^*)\setminus\mathcal{V}$，根据定理 2.6.14，存在一个线性泛函 $L:L_1(P^*)\to\mathbf{R}$，满足：

$$L(V)=0,\quad 对所有 V\in\mathcal{V},\quad 且 L(C)>0.$$

其中，L 具有如下的形式：

$$L(X)=\int XY\mathrm{d}P^*=E^*(XY),\quad X\in L_1(P^*)\quad (对某个 Y\in L_\infty(P^*)).$$

现在可以定义两个等价鞅测度从而得到所需的价格 π_-，π_+. 令：

$$\frac{\mathrm{d}P^+}{\mathrm{d}P^*}(\omega)=1+Y(\omega),\quad \frac{\mathrm{d}P^-}{\mathrm{d}P^*}(\omega)=1-Y(\omega),\quad \omega\in\Omega.$$

前面及以后都假定 Y 满足 $\|Y\|_\infty\leqslant\frac{1}{2}$，则两个测度的 P-密度在 Ω 上都是严格正的，且 P^+，P^-都与 P 等价. 再注意到

$$\int X\mathrm{d}P^{+/-}=E^*(X(1\pm Y))=E^*(X)\pm L(X)=E^*(X)$$

对所有 $X\in\mathcal{V}$成立. 因为 $X=1_\Omega\in\mathcal{V}$，所以 $P^{+/-}(\Omega)=E(1_\Omega)=1$，这就证明了 P^+，P^-是概率测度. 同理，对每一个 i 有

$$E^+(S_{1i})=E^-(S_{1i})=E^*(S_{1i})=S_{0i}(1+r),$$

于是 P^+，P^- 均为等价鞅测度. 在测度 P^+，P^- 下，对应的公平价格分别为

$$\pi_+ = E^+(C^*) = E^*(C^*) + \frac{L(C)}{1+r},$$

$$\pi_- = E^-(C^*) = E^*(C^*) - \frac{L(C)}{1+r}.$$

因为 $L(C)>0$，结论得证. ■

定理 2.6.15 有如下推论.

推论 2.6.16 如果未定权益 C 满足 $E^*(C)<\infty$，并且无套利价格集合不是开区间，则 C 是可复制的.

2.7 一般情形下鞅测度的构造

前面已经知道，可能存在多个不同的等价鞅测度，本节进一步讨论如何求出这些测度. 在有限模型中，可以通过解线性方程组求得，如例 2.5.8. 在一般情形下，例如，在 $\Omega=\mathbf{R}$ 且 $S_1(\omega)=\omega$ 时，$\mathcal{F}=\mathcal{B}(\mathbf{R})$ 上的实际概率测度 P 与时刻 1 时的股价 S_1 有关，可以由分布函数 F 求得. 下列建立在 Esscher 转换上的构造方法是一种常用方法.

对应的数学问题为：要构造一个概率测度 P^* 使得

70

$$E^*(S_1^*) = S_0 \quad \Leftrightarrow \quad E^*(\Delta S) = 0,$$

其中 $\Delta S = S_1^* - S_0$. 构造的基本思想如下：首先通过下列 P-密度

$$f_a(\omega) = \frac{\mathrm{d}P^*}{\mathrm{d}P}(\omega) = \frac{\exp(a\omega)}{E\exp(a\Delta S)}, \quad \omega \in \Omega$$

来定义测度 P^*，其中 a 为参数. 则可得到

$$E^*(\Delta S) = E\left(\frac{\Delta S\exp(a\Delta S)}{E[\exp(a\Delta S)]}\right) = \frac{E(\Delta S\exp(a\Delta S))}{E\exp(a\Delta S)}.$$

上式右边可以写为 $h'(a)/h(a)$，其中指数函数

$$h(a) = E\exp(a\Delta S) = \int \exp(a\Delta S)\mathrm{d}P, \quad a \in \mathbf{R}.$$

事实上，只要 $h(a)<\infty$，则可以在该积分号下求微分. 因此，对 h 的极小值点 a^*，均有 $E^*(\Delta S)=0$. 所以，P-密度 $f_{a^*}(\omega)=\exp(a^*\omega)/E\exp(a^*\Delta S)$，$\omega\in\Omega$ 对应的概率测度 P^* 满足：$P^*(A)=\int_A f_{a^*}(\omega)\mathrm{d}P(\omega)$，$P^*$ 就是所求的一个等价鞅测度. 注意到，如果知道 $\Delta S = S_1^* - S_0$ 或 S_1 在实际测度 P 下的分布，那么 $h(a)$ 和 a^* 都可以得到. 实际上，如果在 P 测度下 $S_1 \sim F$，则有 $\Delta S = S_1/(1+r) - S_0 \sim H(x) = F((1+r)(x-S_0))$，于是得到

$$h(a) = \int \exp(ax)\mathrm{d}H(x),$$

上述构造过程要求 $E(\exp(ax))$ 在 P 下存在，否则可以先作一个简单变换. 下面的定理给出一般情况下的结论，且给出了详细的证明.

定理 2.7.1 假设 $P(S_1>S_0)>0$，定义概率测度 Q 为

$$\mathrm{d}Q(\omega) = c\exp(-\omega^2)\mathrm{d}P,$$

其中 $c=(E_P(\exp(-(\Delta S)^2)))^{-1}$. 设

$$h(a)=E_Q(\exp(a\Delta S)),\quad a\in\mathbf{R},$$

为 ΔS 在 Q 下的矩母函数，则 $h(a)<\infty$ 对任意 $a\in\mathbf{R}$ 成立，并且存在 a^* 满足：

$$h(a^*)=h^*:=\inf\{h(a):a\in\mathbf{R}\}.$$

再定义概率测度 P^* 为

$$\mathrm{d}P^*(\omega)=\frac{\exp(a^*\omega)}{h(a^*)}\mathrm{d}Q(\omega),\tag{2.7}$$

则有 $P^*\sim Q\sim P$，且

$$E^*(\Delta S)=E_{P^*}(\Delta S)=0,$$

71

即 P^* 是 P 的等价鞅测度.

证明 显然有 $0<E(\exp[-(\Delta S)^2])\leqslant 1$，且根据 c 的定义，有 $c\in[1,\infty)$，且

$$Q(\Omega)=c\int_\Omega \exp(-x^2)\mathrm{d}P(x)=1$$

因为 Q 由严格正的密度 $\omega\mapsto c\exp(-\omega^2)$ 定义，所以 $Q\sim P$. 又因为函数 $x\mapsto -x^2+ax$ 在 $x=a/2$ 处达到最大值，所以在测度 Q 下，对任意 $a\in\mathbf{R}$，有

$$h(a)=E_Q(\exp(a\Delta S))=cE_P(\exp[-(\Delta S)^2+a\Delta S])\leqslant c\exp(a^2/4)<\infty,$$

所以，$h^*=\inf\{h(a):a\in\mathbf{R}\}<\infty$. 接下来，由参数 $a\in\mathbf{R}$ 定义如下备选概率测度

$$\frac{\mathrm{d}P_a^*}{\mathrm{d}Q}(\omega)=\frac{\exp(a\omega)}{h(a)},\quad \omega\in\Omega,$$

注意到，$\frac{\mathrm{d}P_a^*}{\mathrm{d}P}\geqslant 0$，所以 $P_a^*\sim Q$. 对任意的备选概率测度 P_a^* 均有 $E^*(\Delta S)=\frac{h'(a)}{h(a)}$. 因此，任一使 $h'(a)=0$ 的 a 都对应一个等价鞅测度. 下面将证明：存在某个最小化的 a^*，即存在 a^*，使得

$$h(a^*)=h^*=\inf\{h(a):a\in\mathbf{R}\}.$$

事实上，如若不然，选取数列 $\{a_n\}$ 使得 $h(a_n)\downarrow h^*$，$n\to\infty$，且 $a_n\to\pm\infty$，否则，可以抽取一个收敛子列 $\{a_m\}$ 使得它的极限 $\bar{a}\in(-\infty,\infty)$，但是由 h 的连续性有 $h(\bar{a})=\lim h(a_m)=h^*$，因此，可以找到另外一个更小的 $\bar{a}$，与假设矛盾，所以 $a_n\to\pm\infty$. 定义 $u_n=|a_n|/a_n$，及

$$u=\lim_{n\to\infty}u_n\in\{-1,+1\}.$$

下面证明 $h(a_n)=E_Q(\exp[a_n\Delta S])\to\infty$，$n\to\infty$. 否则，如果 a_n 收敛，$h(a_n)\downarrow h^*<\infty$，在集合 $A_n=\{a_n\Delta S>\delta|a_n|\}$ 上有 $\exp(a_n\Delta S)\geqslant\exp(\delta|a_n|)$. 又因为

$$h(a_n)\geqslant E_Q(\exp(a_n\Delta S)1_{A_n}),$$

所以

$$h(a_n)\geqslant E_Q(\exp[\delta|a_n|]1_{A_n})=\exp(\delta|a_n|)Q(A_n).$$

若 $Q(A_n)$ 一致大于 0，则上式右边趋于无穷. 但是当 $u_n\to u$，且取 $\delta>0$ 使得 δ/u 是 ΔS 的分布函数 F 的连续点时，

$$Q(A_n)=Q(\Delta S>\delta u_n)\to\varepsilon,\quad n\to\infty.$$

这是由于：因为 $P(\Delta S>0)>0 \Leftrightarrow Q(\Delta S>0)>0$ 排除了“对任意的 $\delta>0$，均有 $Q(u\Delta S>\delta)=0$”. 因此，对任意的 $\varepsilon>0$，均可以找到某个 $\delta>0$ 使得 $Q(u\Delta S>\delta)=\varepsilon>0$，其中 δ 可以取为 F 的连续点. ■

下面是一个经典的例子，有助于理解上面的构造过程.

▶**例** 2.7.2 令 $Y \sim N(0, 1)$，定义概率测度 P_λ 使得

72

$$\frac{\mathrm{d}P_\lambda}{\mathrm{d}P}=\frac{\exp(\lambda Y)}{E_P\exp(\lambda Y)}.$$

因为分母是 Y 的矩母函数，等于 $\exp(\lambda^2/2)$. 为了简洁，对任意一个随机变量 Z，记 $E_\lambda(Z)=\int Z\mathrm{d}P_\lambda$，则对任意的可测集 A，有：

$$\begin{aligned}
P_\lambda(Y \in A) &= E_\lambda(\mathbf{1}(Y \in A)) \\
&= \int \mathbf{1}(Y \in A)\frac{\mathrm{d}P_\lambda}{\mathrm{d}P}\mathrm{d}P \\
&= \int \mathbf{1}(Y \in A)\frac{\exp(\lambda Y)}{\exp(\lambda^2/2)}\mathrm{d}P \\
&= \int \mathbf{1}(y \in A)\frac{\exp(\lambda y)}{\exp(\lambda^2/2)}\mathrm{d}P_Y(y) \\
&= \int_A \frac{1}{\sqrt{2\pi}}\exp\left(-\frac{1}{2}(y^2+2\lambda y-\lambda^2)\right)\mathrm{d}y \\
&= \int_A \frac{1}{\sqrt{2\pi}}\exp\left(-\frac{1}{2}(y-\lambda)^2\right)\mathrm{d}y.
\end{aligned}$$

因此，Y 在 P_λ 下服从 $N(\lambda, 1)$分布. ◀

2.8 完备金融市场

定义 2.8.1 如果任一未定权益 $C: \Omega \to \mathbf{R}_+$ 均是可复制的，则称金融市场是**完备的**.

▶**例** 2.8.2 考虑只有一项风险资产和一个银行账户的二叉树模型. 回顾前面的内容，记概率空间为 $\Omega=\{+, -\}$，则任一未定权益 $C: \Omega \to R_+$ 可由 $c_+=C(+)$，$c_-=C(-)$唯一确定. C 是可复制的当且仅当下列方程有解：

$$C(\omega)=\varphi' S_1(\omega)=\varphi_0(1+r)+\varphi_1 S_{11}(\omega), \quad \omega \in \Omega,$$

其中 $\varphi=(\varphi_0, \varphi_1)'$是复制的资产组合. 因此，只需考虑如下两个方程

$$c_+=\varphi_0(1+r)+\varphi_1 s_+, \quad c_-=\varphi_0(1+r)+\varphi_1 s_-,$$

其中 $s_+=S_{11}(+)$，$s_-=S_{11}(-)$. 显然，当 $s_+-s_->0$ 时，方程的唯一解为

$$\varphi_1=\frac{c_+-c_-}{s_+-s_-}, \quad \varphi_0=\frac{c_-s_+-c_+s_-}{(s_+-s_-)(1+r)}$$ ◀

现考虑有限概率空间 $\Omega=(\omega_1, \cdots, \omega_n)(n \in \mathbf{N})$的情形，回忆前面定义过的$((d+1)\times n)$维支付矩阵

73

$$\boldsymbol{D}=\begin{bmatrix} S_{10}(\omega_1), & \cdots, & S_{10}(\omega_n) \\ \vdots & & \vdots \\ S_{1d}(\omega_1), & \cdots, & S_{1d}(\omega_n) \end{bmatrix}.$$

对有限空间 Ω，可以用 n 维向量 $(C(\omega_1), \cdots, C(\omega_n))'$ 来表示未定权益 C，仍记为 C. 未定权益 C 是可复制的当且仅当存在某个组合 $\varphi \in \mathbf{R}^{d+1}$ 使得

$$\boldsymbol{D}'\varphi = C.$$

成立.

显然，在一个完备市场中，对右边的任意一个 C，该线性方程组必定对应一个解 φ. 由线性代数的经典结论，可以得到下列命题.

命题 2.8.3 设概率空间为 $\Omega = (\omega_1, \cdots, \omega_n)$，$n \in \mathbf{N}$. 金融市场是完全的当且仅当支付矩阵 $\boldsymbol{D}$ 的 n 个不相关的行，即矩阵 $\boldsymbol{D}'$ 的秩等于状态的个数 n(行满秩的).

类似地，如果概率空间 Ω 为一般情形(非有限空间)，则情况要复杂得多. 为此，首先回忆一下所有可复制的支付构成的集合

$$\mathcal{V} = \{\varphi' S_1 : \varphi \in \mathbf{R}^{d+1}\},$$

它是一个线性空间. 显然，$\varphi' S_1$ 是 $\sigma(S_{10}, \cdots, S_{1d})$ 可测的随机变量，且

$$E^* |\varphi' S_1| \leqslant \sum_{i=0}^{d} |\varphi_i| E^*(S_{1i}) = \sum_{i=0}^{d} |\varphi_i| S_{0i}(1+r) < \infty$$

对所有 $P^* \in \mathcal{P}$ 成立. 因此 $\mathcal{V} \subset L^1(\Omega, \sigma(S_{10}, \cdots, S_{1d}), P^*)$. 又因为 $P^*(0 \leqslant S_{1i} < \infty) = 1$，所以 $S_{1i} \in L_0(\Omega, \mathcal{F}, P^*)$ 对任意 $i = 0, 1, \cdots, d$ 和 P^* 成立，并且

$$\mathcal{V} \subset L_1(\Omega, \sigma(S_{11}, \cdots, S_{1d}), P^*) \subset L_0(\Omega, \mathcal{F}, P^*) = L_0(\Omega, \mathcal{F}, P),$$

其中，最后一个等式由 $P^* \sim P$ 导出.

引理 2.8.4 对完全金融市场，有

$$\mathcal{V} = L_1(\Omega, \sigma(S_{11}, \cdots, S_{1d}), P^*) = L_0(\Omega, \mathcal{F}, P^*) = L_0(\Omega, \mathcal{F}, P).$$

证明 设 $X \in L_0 = L_0(\Omega, \mathcal{F}, P)$，可以找到未定权益 $C_1, C_2 \in L_0(\Omega, \mathcal{F}, P)$ 使得 $X = C_1 - C_2$. 根据假设，C_1 和 C_2 是可复制的，则存在 $\varphi_1, \varphi_2 \in \mathbf{R}^{d+1}$ 使得 $C_i = \varphi_i' S_1$，$i = 1, 2$. 于是 $X = (\varphi_1 - \varphi_2)' S_1 \in \mathcal{V}$. ■

下面考虑空间 L_p 的维数，令 $n \in \mathbf{N}$ 为某个固定的正整数，对 Ω 作如下分割：将 Ω 分割为非空不交的集合 $A_1, \cdots, A_n$ 的并，其中 $P(A_i) > 0$，$i = 1, \cdots, n$，即

$$\Omega = \bigcup_{i=1}^{n} A_i, \quad P(A_i) > 0, \quad A_i \cap A_j = \varnothing (i \neq j). \tag{2.8}$$

则示性函数 $1_{A_1}, \cdots, 1_{A_n}$ 在 L_p 空间中是线性无关的. 事实上，若

$$\alpha_1 \mathbf{1}_{A_1}(\omega) + \cdots + \alpha_n \mathbf{1}_{A_n}(\omega) = 0, \quad \forall \omega \in \Omega,$$

74

显然有 $\alpha_1 = \cdots = \alpha_n = 0$. 因此，$\dim(L_p) \geqslant n$. 令

$$N = \sup\{n \in \mathbf{N}: 存在满足(2.8)的集合 A_1, \cdots, A_n \in \mathcal{F}\}.$$

如果 $N = \infty$，则规定 L_p 的维数为无穷大. 现设 $N < \infty$，令 $A_1, \cdots, A_N$ 为满足方程(2.8)的分割，则任意一个满足 $A \in \{A_1, \cdots, A_N\}$ 的集合 A 都是原子集，即：对任意满足 $B \subset A$ 的集合 $B \in \mathcal{F}$，或 $P(B) = 0$，或 $P(B) = P(A)$. 否则，可以用 $B = A \cap B$ 去替代 A，并产生一个新的集合 $A \setminus B$，因此得到一个 $N+1$ 个集合的分割，这与假设矛盾. 所以，$A_1, \cdots, A_N$ 是满足条件(2.8)的最大分划. 于是，任意一个随机变量 $X \in L_p$ 在 A_i 上均几乎必然成为常量. 令 $a_i = X(\omega_i)$，$i = 1, \cdots, N$，其中，$\omega_i \in A_i$ 是任意选取的样本点，可得以下结论

$$X = \sum_{i=1}^{N} a_i \mathbf{1}_{A_i}, \quad P\text{-a. s.}$$

这表明示性函数 $\mathbf{1}_{A_1}$，…，$\mathbf{1}_{A_n}$ 构成了向量空间 L_p 的一组基.

故有下列结论.

引理 2.8.5 $L_p = L_p(\Omega, \mathcal{F}, P)$ 是有限维的当且仅当 Ω 存在一个最大的原子集分划 A_1，…，A_N，使得 $P(A_i) > 0$，$i = 1$，…，N 成立. ■

最后给出本节主要结论的一个总结.

定理 2.8.6 在一个无套利市场中，下列结论成立：

(ⅰ) 如果市场是完备的，则存在唯一的等价鞅测度，即 $|\mathcal{P}| = 1$.

(ⅱ) 如果市场存在唯一的等价鞅测度，则 $\dim \mathcal{V} = \dim L_0 \leqslant d + 1$.

证明

(ⅰ) 如果市场是无套利的，且是完备的. 对固定的 $A \in \mathcal{F}$，则未定权益 $C = \mathbf{1}_A$ 是可复制的. 由定理 2.6.4，无套利价格 $\varphi_A' S_0$ 是唯一的，其中 φ_A 为 C 的任意一个复制资产组合. 因此，映射

$$P^* \mapsto P^*(A) = E^*(C^*) = \varphi_A' S_0, \quad P^* \in \mathcal{P},$$

不依赖于 $P^* \in \mathcal{P}$. 由该结论对所有的集合 $A \in \mathcal{F}$ 均成立，所以，$\mathcal{P}$ 有且仅有一个元素.

(ⅱ) 设 $\mathcal{P} = \{P^*\}$. C 是有界的未定权益，即 $C \in L_0(\Omega, \mathcal{F}, P)$，它唯一的无套利价格为 $E^*(C^*) < \infty$. 由推论 2.6.16 知，C 是可复制的，所以市场是完备的. 从而 $\dim \mathcal{V} = \dim L_0$，且 $\dim \mathcal{V} \leqslant d + 1$. 因此，空间 $(\Omega, \mathcal{F}, P)$ 最多具有 $d+1$ 个原子集 A_1，…，A_{d+1}.

75

上述定理对无套利的完备金融市场做了一个很好的描述：任何一个未定权益都具有以下的表达式：

$$C = \sum_{i=1}^{N} c_i \mathbf{1}_{A_i},$$

其中，$c_i \in (0, \infty)(i = 1, \cdots, N)$，$A_1$，…，$A_N \in \mathcal{F}$ 均为原子. 注意到，$c_i = \infty$ 是不可能出现的，否则，结合 $P(A_i) > 0$（从而 $P^*(A) > 0$，对所有 i）得到 $E^*(C) = \infty$. 如果无套利市场上存在比上式更为一般性的未定权益，则市场是不完备的. ■

将本章结论综述如下：

- 如果存在等价鞅测度，则市场上任何可复制的组合的价格都是唯一的.
- 金融市场是无套利的，当且仅当存在一个等价的鞅测度.
- 无套利金融市场是完备的，当且仅当存在唯一的等价鞅测度.
- 在一个完备的无套利金融市场中，任何一个满足 $E(C) < \infty$ 的未定权益 C 均能被对冲.

2.9 评注与延伸阅读

本章内容主要参考了 Föllmer 和 Schied(2004)，Pliska(1997)和 Shiryaev(1999)的相关资料. 任意概率空间 $(\Omega, \mathcal{F}, P)$ 中资产无套利定价的基本定理有多种证法，定理 2.5.6 是

Dalang-Morton-Willinger 定理中 $T=1$ 的特例(见 Dalang 等(1990)),有限样本的定理 2.5.4 参考了 Harrison 和 Pliska(1981). Delbaen 和 Schachermayer(2006)全面地讨论了这些定理,并介绍了相关的数学理论. 要进一步了解 Esscher 变换下的鞅测度构造,可以参考 Shiryaev(1999).

参考文献

Dalang R.C., Morton A. and Willinger W. (1990) Equivalent martingale measures and no-arbitrage in stochastic securities market models. *Stochastics Stochastics Rep.* **29**(2), 185–201.

Delbaen F. and Schachermayer W. (2006) *The Mathematics of Arbitrage*. Springer Finance. Springer-Verlag, Berlin.

Föllmer H. and Schied A. (2004) *Stochastic Finance: An Introduction in Discrete Time*. vol. 27 of *de Gruyter Studies in Mathematics* extended edn. Walter de Gruyter & Co., Berlin.

Harrison J.M. and Pliska S.R. (1981) Martingales and stochastic integrals in the theory of continuous trading. *Stochastic Process. Appl.* **11**(3), 215–260.

Pliska S. (1997) *Introduction to Mathematical Finance*. Blackwell Publishing, Oxford.

Shiryaev A.N. (1999) *Essentials of Stochastic Finance*. vol. 3 of *Advanced Series on Statistical Science & Applied Probability*. World Scientific Publishing Co. Inc., River Edge, NJ. Facts, models, theory, Translated from the Russian manuscript by N. Kruzhilin.

76 ~ 77

第3章　离散时间的金融模型

金融市场上的数据一般为某个时刻观测到的价格、收益和利率等数值或向量，而其中一些以函数形式表出，如收益曲线和价格模型. 人们观测的时间节点可以是离散的、随机的，甚至可以有不同的时间跨度，如年度、季度，或者是每日的，甚至当日的若干数据. 因此，当人们专注于某个金融目标，比如每隔 Δ 秒观测一次股票的交易价格，就得到了一列数或向量 X_1，X_2，…，这意味着人们必须涉及一个时间序列或者离散时间的随机过程.

除了某些特殊情形外，没有理由认为生成时间序列的随机变量是相互独立的. 在非独立随机变量情况下讨论问题的一种方法是引入所谓的鞅差概念，鞅差有不相关的性质. 鞅是数理金融的一个基础概念，在金融统计中有着非常重要的作用，许多经济变量和经济过程都与它有着密切的关系，本章的开头将讨论这部分内容. 本章也引入了鞅的许多重要结论，它们将在后面的各章中被用到.

鞅差序列虽然是不相关的(仅指线性相关)，但当前时刻 n 所对应的值 X_n 与过去的观测值之间却有可能存在某种其他关系，比如，与滞后值 X_{n-1}，X_{n-2}，…之间存在函数关系 $m(X_{n-1}, X_{n-2}, \cdots)$. 如果能用历史数据推断出函数 m，就可以对 X_n 作近乎完美的估计推断. 但在实际金融市场中这经常是做不到的，因为在实际市场中，大量的机构投资者、投机者和中小个人投资者相互影响，给市场造成了随机和噪声干扰；大宗交易、意料之外的经济和政治新闻、新的政策或法规、不可预报的市场声明和警告、对金融市场信息的捕风捉影以及与此相关的一连串连锁反应等，都会对市场的股价造成或大或小的冲击. 值得庆幸的是，以上效应均可通过随机变量来刻画.

79

如果滞后值 X_{n-1}，X_{n-2}，…和 X_n 之间存在一个线性关系，那么可以建立下列模型

$$X_n = \alpha_1 X_{n-1} + \cdots + \alpha_p X_{n-p} + \sigma\varepsilon_n,$$

其中随机变量 ε_n 具有零均值和单位方差，系数 $\alpha_1, \cdots, \alpha_p \in \mathbf{R}$ 是未知待定参数，常数 $\sigma > 0$. ε_n 刻画了市场的随机因素，称为**随机扰动**. 该模型称为 p 阶自回归模型，它是描述历史观测值之间关系的一个基本时间序列模型. 本章也将讨论 ARMA(p, q)模型，甚至更一般的线性过程，它不仅构建了参数模型的一般框架，而且也朝时间序列的非参数估计迈出了重要一步，在非参数估计中，变量之间的关系式是未知的.

上面的自回归方程在扰动项满足某些假设的条件下是有解的，特别地，当扰动项满足序列弱平稳时可解. 由弱平稳知，序列的一阶与二阶矩均不随时间变化，即弱平稳序列的均值和方差关于时间都是不变的. (弱)平稳性条件是离散时间模型的关键性条件.

图 3-1 分别给出了 4 幅样本数为 500 的时间序列模拟图表，该图表明，平稳序列可以有非常不同的路径，其中有的路径中大量的连续观测值紧紧地挨在一起，有的路径则波动

80

较大，形成具有不同波动率的观测簇. 使用 GARCH 模型可以较好地描述带条件波动率的时间序列模型.

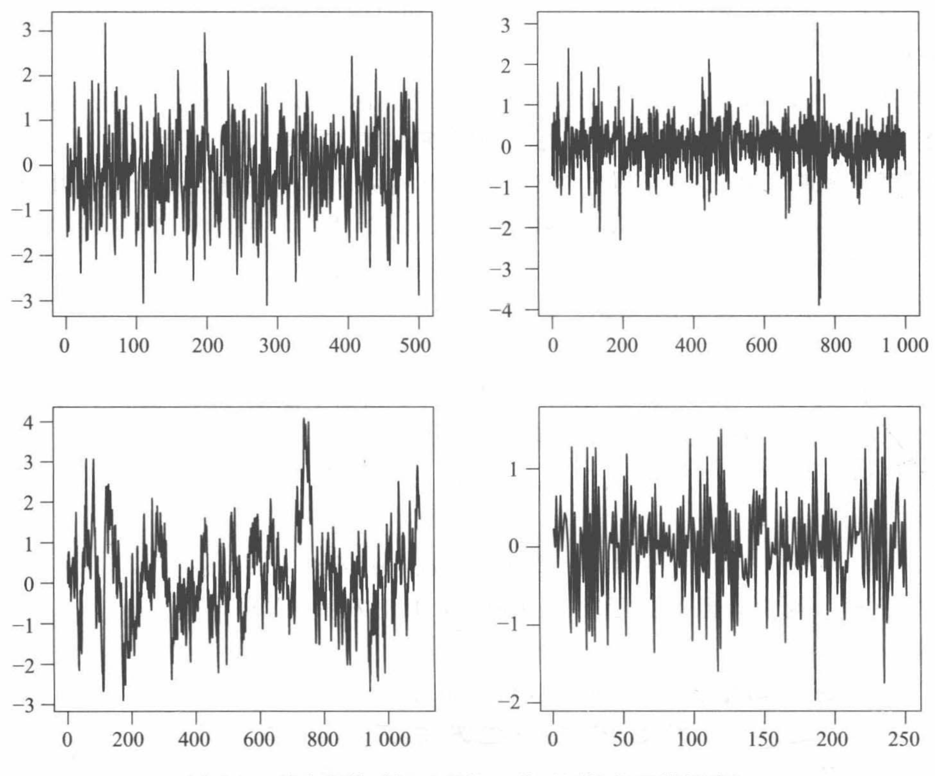


图 3-1 平稳离散时间过程的 4 种不同的典型模拟路径

如果觉得上面的参数模型不可信，那么应使用下列更为接近实际情况的非参数模型

$$X_n = m(X_{n-1}, X_{n-2}, \cdots, X_{n-p}) + \sigma(X_{n-1}, X_{n-2}, \cdots, X_{n-p})\varepsilon_n, \tag{3.1}$$

其中，m 和 σ 均为待定的非线性函数. 非参数方法建立在许多假定基础上，如假设 m 和 σ 都是光滑函数，本书将在第 8 章和第 9 章讨论这一方法.

3.1 离散时间的随机适应过程

本节首先介绍时间序列的一些概念，它们将在后面的章节中用到. 滤子是描述随机变量或向量序列所包含的信息流的一个数学工具，而适应过程是指时间序列模型不依赖于未来信息流的性质. 在给出这些概念之前，先引入时间序列的严格定义.

定义 3.1.1 **一个离散时间的随机过程**是一个随机变量序列 $\{X_n : n \geqslant 0\}$，其中

$$X_n : (\Omega, \mathcal{F}, P) \to (\mathbf{R}, \mathcal{B}), \quad n \geqslant 0.$$

下面介绍滤子以及关于滤子适应的随机过程的概念，滤子在概率论中用于描述信息流. 设$(\Omega, \mathcal{F}, P)$为一个概率空间，$\mathcal{B}$为 $\mathbf{R}$ 的所有 Borel 集构成的集合，如果所有的$\{X \in$

$A\}$(其中 $A\in\mathcal{B}$)均是$\mathcal{F}$的元素，那么称函数 X：$\Omega\to\mathbf{R}$ 是**可测的**. 若离散随机变量的取值为 x_1，x_2，…，则该可测条件就等价于$\{X=x_i\}\in\mathcal{F}$，$i=1$，2，….

考虑只有一个资产的两期模型，其价格在每期对应的上升(下降)因子均为 u(或 d)，则可以设 $\Omega=\{\omega_1, \omega_2, \omega_3, \omega_4\}$，其中 $\omega_1=\{+,+\}$，$\omega_2=\{+,-\}$，$\omega_3=\{-,+\}$，$\omega_4=\{-,-\}$且$\mathcal{F}_2:=\mathcal{F}=\mathrm{Pot}(\Omega)$. 而 $A=\{\omega_1, \omega_2\}$就表示股价在第一期上升的事件. 记$\mathcal{F}_0=\{\varnothing, \Omega\}$，$\mathcal{F}_1=\sigma(A)=\{\varnothing, A, A^c, \Omega\}$，则有$\mathcal{F}_0\subset\mathcal{F}_1\subset\mathcal{F}_2\subset\mathcal{F}$.

为了更好地理解关于$\mathcal{F}_1$可测的含义，考虑下列两个随机变量 X 和 Y，其中 X 关于$\mathcal{F}_1$可测，而 Y 关于$\mathcal{F}_1$不可测. 设 X:$\Omega\to\mathbf{R}$ 满足 $X(\omega_1)=X(\omega_2)=100$，$X(\omega_3)=X(\omega_4)=50$，注意到 X 在$\mathcal{F}_1$的元素集上取值为常数，X 的结果仅仅依赖于截至 $t=1$ 时刻时已经发生的事件，而该信息流由 σ 代数$\mathcal{F}_1$给出. 具体地有

81

$$\{X=100\}=\{\omega_1,\omega_2\}=A\in\mathcal{F}_1,\quad \{X=50\}=\{\omega_3,\omega_4\}=A^c\in\mathcal{F}_1$$

成立，从而 X 关于$\mathcal{F}_1$可测，并且有$\mathcal{F}_1=\sigma(X)$. $\mathcal{F}_2$给出了截至时刻 2 的更多的信息，整个过程的信息流由$\{\varnothing, \Omega\}\subset\mathcal{F}_1\subset\mathcal{F}_2=\mathcal{F}$给出.

现考虑随机变量 Y，其定义为 $Y(\omega_1)=200$，$Y(\omega_2)=150$，$Y(\omega_3)=150$，$Y(\omega_4)=100$，因为

$$\{Y=200\}=\{\omega_1\}\notin\mathcal{F}_1,$$

所以 Y 关于$\mathcal{F}_1$不可测，但 Y 关于$\mathcal{F}_2$是可测的，且有$\mathcal{F}_2=\sigma(Y)$.

定义 3.1.2

(ⅰ) $(\Omega, \mathcal{F})$上的**滤子**是下列的子 σ 代数序列

$$\mathcal{F}_0\subset\mathcal{F}_1\subset\cdots\subset\mathcal{F}_n\subset\mathcal{F},\quad n\in\mathbf{N}.$$

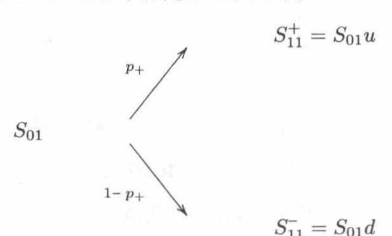
(ⅱ) 一个随机过程$\{X_n:n\geqslant 0\}$称为$\mathcal{F}_n$适应(或适应于$\{\mathcal{F}_n\}$)的，如果对任意的 $n\in\mathbf{N}_0$，均有 X_n 是$\mathcal{F}_n$可测的.

对于给定的过程$\{X_n\}$，通常考虑下列的**自然滤子**

$$\mathcal{F}_0=\{\varnothing,\Omega\},\quad \mathcal{F}_n=\sigma(X_0,\cdots,X_n),\quad n\in\mathbf{N}_0.$$

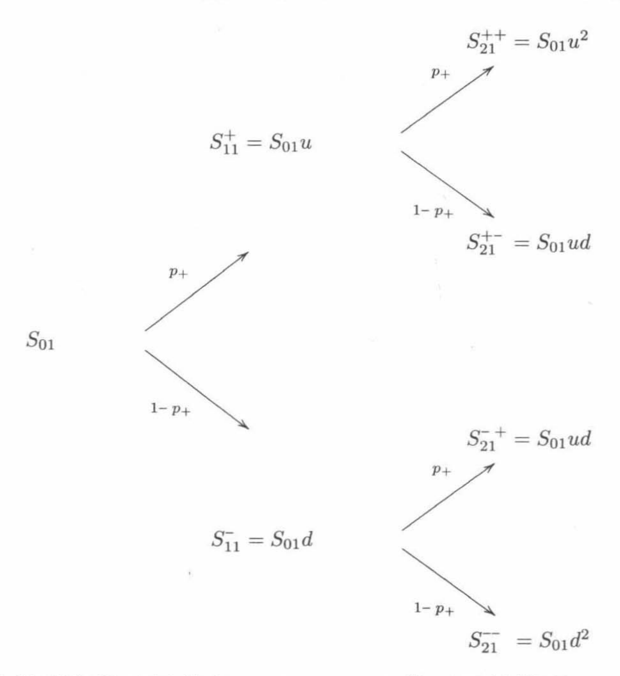
后面如果没有特别指明，本书的滤子均指自然滤子.

下一个需要用到的概念是随机变量在给定 σ 代数下的条件期望，其定义和性质请参考附录 A.3. 为了方便读者，这里在只有无风险银行账户和单只股票的二叉树模型框架下给出条件期望的概念和基本性质，为了能推广到有 d 只股票的情形，用 S_{t1} 表示 t 时刻时的股票价格，给定因子 u 和 d，模型可以用单期模型表示为



82

对单期模型产生的两个叶节点再次扩展一个期，便得到下列二叉树



不断重复这个过程就得到离散时间节点 $t=1$，…，T 的二叉树模型.

注意到 S_{21} 的可能取值为 $\{S_{21}^{++}, S_{21}^{+-}, S_{21}^{--}\}$，对应的概率集为 $\{p_+^2, 2p_+(1-p_+), (1-p_+)^2\}$，它表示股价在第一期的结果已经知道的条件下，第二期将可能会发生的变化. 它的期望为

$$E(S_{21})=S_{21}^{++}p_+^2+2S_{21}^{+-}p_+(1-p_+)+S_{21}^{--}(1-p_+)^2.$$

设

$$\mathcal{F}_0=\{\varnothing,\Omega\},\quad \mathcal{F}_1=\sigma(S_{11}),\quad \mathcal{F}_2=\sigma(S_{11},S_{22}).$$

则$\mathcal{F}_0\subset\mathcal{F}_1\subset\mathcal{F}_2\subset\mathcal{F}$，$\mathcal{F}_t$ 表示在 t 时刻能够得到的信息. 通常，某个节点的条件期望可以用上面的方法来计算，它是一个 S_{11} 的函数，即一个关于$\mathcal{F}_1$ 可测的随机变量. 因为事件 $A\in\mathcal{F}_1$ 仅与第一个坐标有关，即为 ω_1 的函数. 若 $\omega_1=+$，即 $\omega=(+,+)$或 $\omega=(+,-)$，则 S_{21} 可能取值为 S_{21}^{++}，S_{21}^{+-}，从而

$$E(S_{21}\mid\mathcal{F}_1)(\omega)=p_+S_{21}^{++}+(1-p_+)S_{21}^{+-}.$$

同理，$\omega_1=-$ 对应 $\omega\in\{(-,+),(-,-)\}$，于是

$$E(S_{21}\mid\mathcal{F}_1)(\omega)=p_+S_{21}^{-+}+(1-p_+)S_{21}^{--}.$$

因此，$E(S_{21}\mid\mathcal{F}_1)$是 Ω 上的随机变量，根据附录中条件期望的计算法则，有

$$\begin{aligned}E(E(S_{21}\mid\mathcal{F}_1))&=E(\mathbf{1}_\Omega E(S_{21}\mid\mathcal{F}_1))\\&=E(\mathbf{1}_\Omega S_{21})\\&=E(S_{21}).\end{aligned}$$

上例表明，即使在最简单的金融市场模型中，也可能出现条件期望.

在本书讨论的两期二叉树模型中，资产在 t 时刻的价格是 $t-1$ 时刻价格的函数，因此它简化了条件概率的计算.

定义 3.1.3 设随机过程$\{X_n:n\in\mathbf{N}\}$配备了自然滤子$\{\mathcal{F}_n:n\in\mathbf{N}\}$，$\mathcal{F}_n=\sigma(X_1,\cdots,X_n)$，$n\in\mathbf{N}$，称$\{X_n:n\in\mathbf{N}\}$为**一个马尔可夫过程**(或马尔可夫链)，如果

$$P(X_{n+1}\in A\mid\mathcal{F}_n)=P(X_{n+1}\in A\mid X_n)$$

对所有的$A\in\mathcal{F}$成立.

由定义可知，若$\{X_n:n\in\mathbf{N}\}$是一个马尔可夫过程，则事件$\{X_{n+1}\in A\}$关于到n时刻为止的信息的条件概率仅与最近的值X_n有关.

▶**例 3.1.4** 设$\{\xi_n:n\in\mathbf{N}\}$为一列 i.i.d. 随机变量，取值空间为$\{-1,1\}$，且$p=P(\xi_1=1)$. 定义$S_0=0$，$S_n=\sum_{i=1}^{n}\xi_i$，$n\in\mathbf{N}$，则$\{S_n\}$为马尔可夫过程，且

$$P(S_n=j\mid S_{n-1}=i)=P(\xi_n=j-i\mid S_{n-1}=i)$$

$$=\begin{cases}p, & i=j-1,\\ 1-p, & i=j+1,\\ 0, & \text{其他情形},\end{cases}$$

其中i，$j\in\{-n,\cdots,n\}$，这个条件概率是i，j的函数，记为$p(i,j)$，称$p(i,j)$为马尔可夫链从初始状态i跳到状态j的**转移概率**. ◀

将上例推广到一般情形，设随机变量序列X_0，…，X_T的取值范围为S，称S为**状态空间**，若对满足$P(X_0=x_0,\cdots,X_{n-1}=x_{n-1})>0$的$x_n\in S$，$n=1$，…，$T$，

$$P(X_n=x_n\mid X_0=x_0,\cdots,X_{n-1}=x_{n-1})$$

均等于

$$p(x_{n-1},x_n)=P(X_n=x_n\mid X_{n-1}=x_{n-1}),$$

则称X_0，…，X_T为一个马尔可夫链. $P(X_0=x_0,\cdots,X_{n-1}=x_{n-1})>0$保证了路径$x_0$，…，$x_{n-1}$是可观测的，称矩阵$\boldsymbol{P}=(p(i,j))_{i,j\in S}$为马尔可夫过程的**转移矩阵**；记$p_i=P(X_0=i)$，$i\in S$，称行向量$p_0=(p_i)_{i\in S}$为**初始分布**. 容易看出该过程一步之后的(边际)分布为

$$p_j^{(1)}=P(X_1=j)=\sum_{i\in S}P(X_1=j\mid X_0=i)P(X_0=i)=(p_0\boldsymbol{P})_{aj},$$

其中$j\in S$，$p^{(1)}=(p_j^{(1)})_{j\in S}$满足$p^{(1)}=p_0\boldsymbol{P}$. 一般地，经过$n$步后，对应的边际分布$p^{(n)}=(P(X_n=j))_{j\in S}$为

84

$$p^{(n)}=p_0\boldsymbol{P}^n,$$

其中$\boldsymbol{P}^n=(p^n(i,j))_{i,j\in S}$称为$\boldsymbol{n}$ **步转移矩阵**. 利用$\boldsymbol{P}^{m+n}=\boldsymbol{P}^m\cdot\boldsymbol{P}^n$可以得到著名的 Chapman-Kolmogorov 方程

$$p^{(m+n)}(x,y)=\sum_{z\in S}p^{(m)}(x,z)p^{(n)}(z,y).$$

3.2 鞅和鞅差序列

下列内容是为引入鞅做准备的，它引导读者对鞅有一个初步的认识. 记t时刻的资产价格为S_t，其增量$Y_t=S_t-S_{t-1}(t=1,\cdots,T)$是 i.i.d.，且

$$P(Y_t=1)=p,\quad P(Y_t=-1)=1-p.$$

显然 $Y_0=0$，设投资者在 $t=0$ 到 $t=1$ 时段持有的资产份额为 φ_1，从 $t-1$ 至 t 时段持有的资产份额为 φ_t. 考虑任意投资策略，其中 φ_t 依赖于历史价格，即

$$\varphi_t = \varphi_t(Y_0,\cdots,Y_{t-1})$$

可能是 Y_1，…，Y_{t-1}(或者加上 S_0)的函数，因此也是 S_0，…，S_{t-1} 的函数. 设$\mathcal{F}_0=\{\varnothing,\Omega\}$，$\mathcal{F}_t=\sigma(Y_1, \cdots, Y_t)$，$t\geqslant 1$，则 φ_t 关于$\mathcal{F}_{t-1}$可测. 注意到，资产的当前价格可以通过下列的递归方式来计算

$$X_t = X_{t-1} + \varphi_t(Y_0,\cdots,Y_{t-1})Y_t,\quad t=1,\cdots,T,\quad X_0=0,$$

即

$$X_t = \sum_{i=1}^{t}\varphi_i(Y_0,\cdots,Y_{i-1})Y_i.$$

下面计算 X_t 关于到 $t-1$ 时刻为止的信息$\mathcal{F}_{t-1}=\sigma(Y_1, \cdots, Y_{t-1})$的条件期望

$$\begin{aligned} E(X_t\,|\,\mathcal{F}_{t-1}) &= E(X_t\,|\,Y_0,\cdots,Y_{t-1}) \\ &= E\Big(\sum_{i=1}^{t}\varphi_i(Y_0,\cdots,Y_{i-1})Y_i\,|\,Y_0,\cdots,Y_{t-1}\Big) \\ &= \sum_{i=1}^{t}\varphi_i(Y_0,\cdots,Y_{i-1})E(Y_i\,|\,Y_0,\cdots,Y_{t-1}), \end{aligned}$$

85

其中用到了条件期望的线性性质，即条件期望内的一个常数因子，可以放在条件期望的外面. 再由 Y_0，…，Y_T 的独立性，有

$$E(Y_i\,|\,Y_0,\cdots,Y_{t-1}) = \begin{cases} Y_i, & i\leqslant t-1, \\ E(Y_i)=2p-1, & i>t-1. \end{cases}$$

因此，资产在 t 时刻的价值关于到 $t-1$ 时刻为止信息的条件期望为

$$E(X_t\,|\,\mathcal{F}_{t-1}) = X_{t-1} + (2p-1)\varphi_t(Y_0,\cdots,Y_{t-1}).$$

需要指出的是：上述所有的计算结果都是在 a.s. 意义下成立的. 特别地，取 $p=1/2$ 得到 $E(X_t\,|\,\mathcal{F}_{t-1})=X_{t-1}$，即任何投资策略都不改变此种情形的条件期望；若 $p>1/2$，同时 $\varphi_t>0$，则有 $E(X_t\,|\,\mathcal{F}_{t-1})>X_{t-1}$，这表明这种情况下投资是明智的. 类似地，若 $p<1/2$，则应持空头.

3.2.0.1 鞅

定义 3.2.1 $\{X_n\}$称为$\mathcal{F}_n$ **鞅**或**关于**$\mathcal{F}_n$ **的鞅**，如果

(ⅰ) $\{X_n\}$是$\mathcal{F}_n$ 适应的，

(ⅱ) $\forall n$，$E|X_n|<\infty$，

(ⅲ) $\forall n$，$E(X_{n+1}\,|\,\mathcal{F}_n)=X_n$，a.s..

如果条件(ⅰ)、(ⅱ)及下列条件(ⅳ)成立，那么称$\{X_n\}$为$\mathcal{F}_n$ **下鞅**，

(ⅳ) $\forall n$，$E(X_{n+1}\,|\,\mathcal{F}_n)\geqslant X_n$，a.s..

如果条件(ⅰ)、(ⅱ)及下列条件(ⅴ)成立，那么称$\{X_n\}$为$\mathcal{F}_n$ **上鞅**，

(ⅴ) $\forall n$，$E(X_{n+1}\,|\,\mathcal{F}_n)\leqslant X_n$，a.s..

注 3.2.2 利用条件期望的全期望公式易证，如果$\{X_n\}$为下鞅，那么

$$E(X_{n+m}\,|\,\mathcal{F}_n) \geqslant X_n$$

对所有 n，$m\geqslant 0$ 均成立．对于上鞅，只需把$\geqslant$替换成$\leqslant$即可；如果$\{X_n\}$是鞅，那么

$$E(X_{n+m}\mid\mathcal{F}_n)=X_n$$

对所有 n，$m\geqslant 0$ 均成立．

注 3.2.3 如果$\{X_n\}$是$\mathcal{F}_n$ 下鞅，$\{\mathcal{G}_n\}$是一个子滤子，即$\mathcal{G}_n\subset\mathcal{F}_n$，$\forall n$，只要$\{X_n\}$关于$\mathcal{G}_n$ 适应，则$\{X_n\}$关于$\mathcal{G}_n$ 也是一个下鞅．事实上，

$$E(X_{n+1}\mid\mathcal{G}_n)=E(E(X_{n+1}\mid\mathcal{F}_n)\mid\mathcal{G}_n)\geqslant E(X_n\mid\mathcal{G}_n)=X_n.$$

86

该结论对鞅和上鞅也成立．

▶**例 3.2.4** 下面是两个经典的例子．

（ⅰ）设 X_1，$X_2\cdots$为相互独立的随机变量，对 $i\geqslant 1$ 有 $E|X_i|<\infty$，$E(X_i)=0$，记

$$S_n=\sum_{i=1}^n X_i,\quad n\in\mathbf{N}.$$

则$\{S_n:n\in\mathbf{N}_0\}$关于滤子$\mathcal{F}_n=\sigma(X_1,\cdots,X_n)$是鞅，事实上，对 $n\geqslant 1$ 有

$$E|S_n|\leqslant\sum_{i=1}^n|E(X_i)|\leqslant\sum_{i=1}^n E|X_i|<\infty,$$

且对 $n\geqslant 0$，有

$$\begin{aligned}E(S_{n+1}\mid\mathcal{F}_n)&=E(S_n+X_{n+1}\mid\mathcal{F}_n)\\&=E(S_n\mid\mathcal{F}_n)+E(X_{n+1}\mid\mathcal{F}_n)\\&=S_n,\end{aligned}$$

上式在几乎必然意义下成立是因为：首先 S_n 关于$\mathcal{F}_n$ 可测，其次因为根据假设 X_{n+1}独立于$\mathcal{F}_n=\sigma(X_1,\cdots,X_n)$，所以 $E(X_{n+1}\mid\mathcal{F}_n)=E(X_{n+1})=0$．

（ⅱ）设 Z_1，$\cdots$，Z_n 是 i. i. d.，且 $E|Z_1|<\infty$，$E(Z_1)=1$．记

$$T_n=\prod_{i=1}^n Z_i,\quad n\geqslant 1.$$

则 T_n 关于滤子$\mathcal{F}_n=\sigma(Z_1,\cdots,Z_n)$，$n\in\mathbf{N}_0$ 是一个鞅．事实上，对任意的 n，由独立性有

$$E|T_n|\leqslant\prod_{i=1}^n E|Z_i|<\infty,$$

又因为

$$\begin{aligned}E(T_{n+1}\mid\mathcal{F}_n)&=E\Big(Z_{n+1}\prod_{i=1}^n Z_i\mid\mathcal{F}_n\Big)\\&=E(Z_{n+1}\mid\mathcal{F}_n)\prod_{i=1}^n Z_i\\&=E(Z_{n+1})T_n\\&=T_n,\end{aligned}$$

87

几乎必然成立，其中用到了 $\prod_{i=1}^n Z_i$ 关于$\mathcal{F}_n$ 可测且 Z_{n+1}关于$\mathcal{F}_n$ 独立的事实． ◀

上面例子中的（ⅱ）在统计学中有着重要的应用，设有下列的随机样本

$$X_1,\cdots,X_n\overset{\text{i.i.d.}}{\sim}f_{\theta_0}(x),$$

其中 f_{θ_0} 属于带参数的密度函数族 $\{f_\theta:\theta\in\Theta\}$，其中 $\Theta\in\mathbf{R}^k(k\in\mathbf{N})$ 表示一个 k 维参数空间. 假设 θ_0 是未知的，但可以通过 X_1，…，X_n 来估计. 定义下列函数

$$L(\theta|X_1,\cdots,X_n)=\prod_{i=1}^n f_\theta(X_i),\quad \theta\in\Theta,$$

称这个函数为**似然函数**，设这个函数关于 θ 存在最大值，则称对应的最大值点 θ_0 为**极大似然估计量**(m. l. e.). 对充分小的 $\Delta x>0$，

$$\int_{[X_i-\Delta x/2,X_i+\Delta x/2]} f_\theta(u)\mathrm{d}u\approx\Delta x f_\theta(X_i)$$

表示在独立重复实验条件下得到的观测值接近于 X_i 的概率. 从这个意义上讲，$L(\theta|X_1,\cdots,X_n)$ 可以认为是取到样本 X_1，…，X_n 的概率的近似值，而 m. l. e. 是在取得样本观测值条件下，使该概率最大的参数值.

极大似然估计也可以由假设检验来求得，设

$$H_0:\theta_0=\theta^{(0)}$$

对应的备择假设为

$$H_1:\theta_0=\theta^{(1)},$$

其中 $\theta^{(1)}$，$\theta^{(0)}$ 均为给定的常数. 定义**似然比统计量**为

$$L_n=\frac{\prod_{i=1}^n f_{\theta^{(1)}}(X_i)}{\prod_{i=1}^n f_{\theta^{(0)}}(X_i)}=\prod_{i=1}^n\frac{f_{\theta^{(1)}}(X_i)}{f_{\theta^{(0)}}(X_i)},$$

该统计量是 $\theta^{(1)}$，$\theta^{(0)}$ 对应的似然函数的比的乘积，若 L_n 较大，则拒绝原假设而接受备择假设.

从例 3.2.4 中的(ⅱ)知，L_n 在原假设成立条件下是一个鞅. 再者，在比率变量 $Z_i=\frac{f_{\theta^{(1)}}(X_i)}{f_{\theta^{(0)}}(X_i)}(i=1,\cdots,n)$ 属于空间 L_1 的条件下，有

$$\begin{aligned}E_{\theta^{(0)}}\left(\frac{f_{\theta^{(1)}}(X_i)}{f_{\theta^{(0)}}(X_i)}\right)&=\int\frac{f_{\theta^{(1)}}(x)}{f_{\theta^{(0)}}(x)}f_{\theta^{(0)}}(x)\mathrm{d}x\\&=\int f_{\theta^{(1)}}(x)\mathrm{d}x=1,\end{aligned}$$

其中，$E_{\theta^{(0)}}$ 表示在 $X_i\sim f_{\theta^{(0)}}(x)$ 下求期望.

下面的例子虽然简单，但是非常重要.

88

▶例 3.2.5(Lévy 鞅) 设 ξ 为一个任意的随机变量，满足 $E|\xi|<\infty$，$\{\mathcal{F}_n:n\in\mathbf{N}_0\}$ 为任意一个滤子，考虑下列条件期望的序列

$$X_n=E(\xi|\mathcal{F}_n),\quad n\in\mathbf{N}_0,$$

它是一个鞅过程，称为 **Lévy 鞅**. 事实上，因为

$$E|X_n|=E|E(\xi|\mathcal{F}_n)|\leqslant E(E(|\xi|)\mathcal{F}_n)=E|\xi|<\infty,$$

所以，对 $\forall n$ 有 $X_n\in L_1$. 其次，由条件期望的全期望公式可得鞅条件成立，即由 $\mathcal{F}_n\subset\mathcal{F}_{n+1}$ 知，对 $n\in\mathbf{N}$ 有

$$\begin{aligned}E(X_{n+1}|\mathcal{F}_n) &= E(E(\xi|\mathcal{F}_{n+1})|\mathcal{F}_n)\\ &= E(\xi|\mathcal{F}_n)\\ &= X_n.\end{aligned}$$
◀

定理 3.2.6

(ⅰ) 如果 X_n 是$\mathcal{F}_n$ 下鞅，那么 $-X_n$ 是$\mathcal{F}_n$ 上鞅.

(ⅱ) 如果 X_n 是$\mathcal{F}_n$ 鞅，$\varphi:\mathbf{R}\to\mathbf{R}$ 是一个凸函数，且 $\forall n$，$E|\varphi(X_n)|<\infty$，那么 $\varphi(X_n)$ 是$\mathcal{F}_n$ 下鞅.

(ⅲ) 若$\{X_n\}$是下鞅，则对任意 $\forall n\leqslant m$，有 $E(X_n)\leqslant E(X_m)$.

证明 (ⅰ)是显然的. (ⅱ)由条件期望的 Jensen 不等式得到，见附录定理 A.3.1(xi). ■

▶**例** 3.2.7 设 X_1，X_2，…为独立随机变量，且对所有 $i\in\mathbf{N}$，$X_i\in L_1$，$E(X_i)=0$，则 $S_n=\sum_{i=1}^{n}X_i$，$n\geqslant1$ 为一个鞅，S_n^2 为一个下鞅，且不等式 $E(S_n^2)\geqslant E(S_{n+1}^2)$对所有的 n 均成立. ◀

3.2.0.2 鞅差序列

鞅差序列在金融统计中起着重要的作用. 在增加一些矩条件下，鞅差序列是一个后面将介绍的白噪声过程，而白噪声过程是鞅在实际中的重要应用，在金融数据模型中白噪声过程常常被用来代替 i.i.d. 序列.

89

定理 3.2.8(鞅差序列(MDS)) 设$\{M_n\}$是一个$\mathcal{F}_n$ 鞅，称随机变量序列

$$\varepsilon_n = M_n - M_{n-1},\quad n\geqslant1$$

为一个**鞅差序列**. 反之，若随机变量序列$\{\varepsilon_n\}$满足

(ⅰ) $E(\varepsilon_n)<\infty$对所有 n 成立，

(ⅱ) $\{\varepsilon_n\}$关于$\mathcal{F}_n$ 适应，

(ⅲ) $E(\varepsilon_n|\mathcal{F}_{n-1})=0$，$P$-a.s. 对所有 n 成立，

则 $M_n=\sum_{i=1}^{n}\varepsilon_i$，$n\geqslant0$ 是一个$\mathcal{F}_n$ 鞅，而$\{\varepsilon_n\}$是一个鞅差序列.

证明 因为 M_n 和 M_{n-1}均关于$\mathcal{F}_n$ 可测，则其差 $\varepsilon_n=M_n-M_{n-1}$亦关于$\mathcal{F}_n$ 可测，且 $E|\varepsilon_n|\leqslant E|M_n|+E|M_{n-1}|<\infty$对所有 n 成立. 最后

$$\begin{aligned}E(\varepsilon_n|\mathcal{F}_{n-1}) &= E(M_n|\mathcal{F}_{n-1}) - E(M_{n-1}|\mathcal{F}_{n-1})\\ &= M_{n-1}-M_{n-1}=0,\end{aligned}$$

a.s. 成立，令$\{\varepsilon_n\}$是一个满足(ⅰ)~(ⅲ)的序列，则 $M_n=\sum_{i=1}^{n}\varepsilon_i$ 是$\mathcal{F}_n$ 可测的，对所有的 n 均有 $E|M_n|\leqslant\sum_{i=1}^{n}E|\varepsilon_i|<\infty$ 成立，且

$$E(M_n|\mathcal{F}_{n-1}) = \sum_{i=1}^{n-1}\varepsilon_i + E(\varepsilon_n|\mathcal{F}_{n-1}) = M_{n-1},\text{a.s.},$$

其中上式用到 $\varepsilon_i(i\leqslant n-1)$关于$\mathcal{F}_{n-1}$可测，从而 $E(\varepsilon_i|\mathcal{F}_{n-1})=\varepsilon_i$(a.s.). ■

鞅差序列是零均值的，如果它的二阶矩存在且有限，那么鞅差序列是不相关的. 事实

上，对任意的 $m<n$，因为$\mathcal{F}_m\subset\mathcal{F}_{n-1}$，且 $E(\varepsilon_n|\mathcal{F}_{n-1})=0$，a. s.，所以

$$E(\varepsilon_n|\mathcal{F}_m)=E(E(\varepsilon_n|\mathcal{F}_{n-1}|\mathcal{F}_m)=0,$$

a. s. 成立，从而 $E(\varepsilon_n)=E(E(\varepsilon_n|\mathcal{F}_m))=0$. 若条件 $E(\varepsilon_n^2)<\infty$，$E(\varepsilon_m^2)<\infty$成立，则有

$$\mathrm{Cov}(\varepsilon_n,\varepsilon_m)=E(\varepsilon_n\varepsilon_m)=E(\varepsilon_m E(\varepsilon_n|\mathcal{F}_m))=0.$$

综上所述，有下列重要结论.

引理 3.2.9 设$\{M_n\}$为一个鞅，其鞅差序列为$\{\varepsilon_n\}$，且 $E(\varepsilon_n^2)<\infty$对所有 n 成立，则

90

（ⅰ）$\{\varepsilon_n\}$有相同的均值，且互不相关；

（ⅱ）$\mathrm{Var}(M_n)=\mathrm{Var}(M_0)+\sum_{i=1}^{n}\mathrm{Var}(\varepsilon_i)$.

3.2.1 鞅变换

设 $S_t(t=0,\cdots,T)$为一个股价过程，且初始股价 S_0 已知，记 $t-1$ 到 t 时刻持有的股票份额为 φ_t，其中 φ_t 关于$\mathcal{F}_{t-1}$可测.

定义 3.2.10（可预报过程） 设$\{X_n\}$是一个离散时间的随机过程，$\{\mathcal{F}_n\}$为一个滤子，若对所有的 n，均有 X_n 是$\mathcal{F}_{n-1}$可测的，则称$\{X_n\}$为一个**可预报**（或**可料**）过程.

由定义知，资产组合的投资策略过程$\{\varphi_t\}$是一个可预报过程. 该投资策略过程在 t 时刻的价值 V_t 满足下列的递归等式

$$V_t=V_{t-1}+\varphi_t(S_t-S_{t-1}),$$

从而有 $V_t=V_0+\sum_{i=1}^{t}\varphi_i(S_i-S_{i-1})$ 对所有 $t=0,\cdots,T$ 均成立.

定义 3.2.11（离散随机积分） 设$\{X_n:n\geqslant0\}$为一个$\mathcal{F}_n$ 适应过程，$\{\varphi_n:n\geqslant1\}$为可预报过程，称

$$I_n=\sum_{i=1}^{n}\varphi_i(X_i-X_{i-1}),\quad n\geqslant0,$$

为**离散随机积分**（设其初值为 I_0），记为 $I_n=\int_0^n\varphi_t\mathrm{d}X_t$.

上述记法是为了与连续情形形式上保持一致.

定理 3.2.12 设 X_n 为$\mathcal{F}_n$ 鞅，φ_n 是$\mathcal{F}_n$ 可测的，且

$$\varphi_n\in L_p,\quad X_n\in L_q,$$

其中，p，$q\in[1,\infty]$，$\frac{1}{p}+\frac{1}{q}=1$，则它们对应的离散随机积分是一个$\mathcal{F}_n$ 鞅. 进一步，对任意常数 X_0，

$$X_0+\int_0^n\varphi_t\mathrm{d}X_t=X_0+\sum_{i=1}^{n}\varphi_i(X_i-X_{i-1}),$$

亦有同样的结论.

91

证明 由 Hölder 不等式，$\forall n$ 有

$$E|\varphi_nX_n|\leqslant\|\varphi_n\|_{L_p}\|X_n\|_{L_q}<\infty,$$

因此

$$E|X_0+I_n|=E\left|X_0+\sum_{i=1}^{n}\varphi_i(X_i-X_{i-1})\right|$$
$$\leqslant E|X_0|+\sum_{i=1}^{n}(E|\varphi_iX_i|+E|\varphi_iX_{i-1}|)<\infty,$$

即对 $\forall n$，有 $X_0+I_n\in L_1$ 成立. 因为 φ_i 关于 $\mathcal{F}_{i-1}$ 可测，所以它关于 $\mathcal{F}_n$ 可测，又因为 X_i 关于 $\mathcal{F}_i$ 可测，从而关于 $\mathcal{F}_n$ 可测，则可得

$$E(X_0+I_{n+1}|\mathcal{F}_n)=E\Big(X_0+\sum_{i=1}^{n+1}\varphi_i(X_i-X_{i-1})|\mathcal{F}_n\Big)$$
$$=X_0+\sum_{i=1}^{n}E(\varphi_i(X_i-X_{i-1})|\mathcal{F}_n)+E(\varphi_{n+1}(X_{n+1}-X_n)|\mathcal{F}_n).$$

注意到对 $i=1, \cdots, n$，均有

$$E(\varphi_i(X_i-X_{i-1})|\mathcal{F}_n)=\varphi_i(X_i-X_{i-1})$$

且

$$\begin{aligned}E(\varphi_{n+1}(X_{n+1}-X_n)|\mathcal{F}_n)&=\varphi_{n+1}E(X_{n+1}-X_n|\mathcal{F}_n)\\&=\varphi_{n+1}[E(X_{n+1}|\mathcal{F}_n)-E(X_n|\mathcal{F}_n)]\\&=0.\end{aligned}$$

所以有

$$E(X_0+I_{n+1}|\mathcal{F}_n)=X_0+\sum_{i=1}^{n}\varphi_i(X_i-X_{i+1})=I_n.$$

P-a. s. 成立，则 X_0+I_n 是鞅. ■

定义 3.2.13(鞅变换) 在定理 3.2.12 的条件下，离散随机积分

$$I_n=\sum_{i=1}^{n}\varphi_i(X_i-X_{i-1}),\quad n\geqslant 0,$$

92 称为一个**鞅变换**.

3.2.2 停时、可选抽样定理和极大不等式

设随机序列 V_1，V_2，…表示交易策略(比如某金融工具中的多头头寸)对应的价值过程. 考虑下列问题：根据某种规则(退出策略)在 N 时刻停止交易(其中 N 取值为自然数)，使得投资组合的价值在 N 后一直为 V_N 不变，则对应的价值过程为

$$V_1,V_2,\cdots,V_N,V_N,\cdots \tag{3.2}$$

设$\{V_n\}$是一个鞅，那么对(3.2)所表示的停时序列，鞅性质仍然成立吗？在没有停时的情况下，序列对所有的 n 均有 $E(V_n)=E(V_1)$成立，是否存在一个停止策略，使得停时序列成为一个下鞅，即有 $E(V_N)\geqslant E(V_1)$？

3.2.2.1 停时

显然，在 n 时刻是停止交易或继续交易依赖于到 n 时刻(其中包括时刻 n 在内)为止已获得的信息，从而是否交易的决定应为所获信息的函数. 由于无法利用未来的信息，所

以，从数学意义上讲，该决定是关于$\mathcal{F}_n$可测的，因此导出下列定义.

定义 3.2.14 设$\{\mathcal{F}_n\}$为定义在概率空间$(\Omega, \mathcal{F}, P)$上的滤子，N是取值于$\mathbf{N}_0$的随机变量，如果对所有的$n\in\mathbf{N}$均有$\{N\leqslant n\}\in\mathcal{F}_n$成立，那么称$N$为**停时**.

▶**例 3.2.15** 下面是几个例子及反例.

(ⅰ) 设V_1，V_2，…为某个复合投资策略(即有买也有卖的策略)的价值过程. 设投资者在其价值达到某个上界值$\overline{v}$就停止他的交易活动，那么得到下列停时

$$N_1 = \inf\{t \in \mathbf{N}: V_t \geqslant \overline{v}\}.$$

如果还规定当价值过程下跌到某个阈值$\underline{v}$以下，投资者对其头寸平仓，那么得到

$$N_2 = \inf\{t \in \mathbf{N}: V_t \geqslant \overline{v} \text{ 或 } V_t < \overline{v}\}.$$

易证N_1，N_2是定义 3.2.14 意义下的$\mathcal{F}_n$停时.

(ⅱ) 设X_1，X_2，…是独立随机变量，且对所有的i，有$E|X_i|<\infty$，令$\mathcal{F}_n=\sigma(X_1, X_2, \cdots, X_n)$，$n\in\mathbf{N}_0$，定义$S_0=0$，$S_n=\sum_{i=1}^{n}X_i$，$n\in\mathbf{N}_0$，则

$$N = \inf\{n \geqslant 0: S_n \geqslant 1\}$$

是一个停时. 事实上，

$$\{N = n\} = \{S_1 < 1, \cdots, S_{n-1} < 1, S_n \geqslant 1) \in \mathcal{F}_n.$$

(ⅲ) 反例，

$$U = \sup\{n \geqslant 0: S_n \geqslant 1\}$$

93

不是停时. 事实上

$\{U = n\} = \{S_1, \cdots, S_{n-1} \in \mathbf{R}, S_n \geqslant 1, S_{n+1} < 1, S_{n+2} < 1, \cdots\}$ 依赖于所有的S_i. ◀

定义 3.2.16 设$\{X_n: n\geqslant 0\}$是一个$\mathcal{F}_n$适应过程，N是一个停时，设$X^N=\{X_n^N: n\geqslant 0\}$，且

$$X_n^N = X_{\min(n,N)} = \begin{cases} X_n, & n \leqslant N, \\ X_N, & n > N, \end{cases}$$

则称X^N为一个**停时过程**.

使用记号$a\wedge b=\min(a, b)$，上述过程可记为$X_n^N=X_{n\wedge N}$.

引理 3.2.17 设$\varphi_n=\mathbf{1}\{n\leqslant N\}$，$n\in\mathbf{N}$，则停时过程可以表示为离散时间的随机积分

$$X_n^N = X_0 + \int_0^n \varphi_t \mathrm{d}X_t = X_0 + \sum_{i=1}^{n} \varphi_i (X_i - X_{i-1}).$$

证明 下列证明过程虽然简单，但富有启发性. 首先

$$\begin{aligned} X_0 + \sum_{i=1}^{n} \varphi_i (X_i - X_{i-1}) &= X_0 + \sum_{i=1}^{n} \mathbf{1}(i \leqslant N)(X_i - X_{i-1}) \\ &= X_0 + \sum_{i=1}^{n\wedge N} (X_i - X_{i-1}) \\ &= \begin{cases} X_n, & n \leqslant N, \\ X_N, & n > N. \end{cases} \end{aligned}$$

现在只需证明φ_n是$\mathcal{F}_{n-1}$可测的即可，事实上，由$\{n\leqslant N\}=\{N\leqslant n-1\}^c\in\mathcal{F}_{n-1}$立即得证. ■

3.2.2.2 可选抽样定理

设 X_n 是关于滤子$\mathcal{F}_n$ 的下鞅(或鞅)，考虑停时列 N_1，N_2，…，设它们几乎必然有序，即不等式 $N_1 \leqslant N_2 \leqslant \cdots$ 以概率 1 成立. 现在考虑一个问题，如果 X_n 是下鞅(鞅)，那么 X_{N_1}，X_{N_2}，…是否仍然为下鞅(鞅)?

94

由鞅变换理论可以得到下列结论.

定理 3.2.18 设$\{X_n\}$为$\mathcal{F}_n$ 鞅，N 为停时，那么停时过程$\{X_n^N\}$也是$\mathcal{F}_n$ 鞅，且有

$$E(X_n^N) = E(X_n) = E(X_1) = X_0.$$

若$\{X_n\}$为$\mathcal{F}_n$ 下鞅(上鞅)，则$\{X_n^N\}$也是$\mathcal{F}_n$ 下鞅(上鞅).

证明 若 $\varphi_n = \mathbf{1}(n \leqslant N)$，则$\{X_n^N\}$是关于$\{X_n\}$的一个鞅变换. 根据$\{\varphi_n\}$的有界性，由定理 3.2.12 可直接推出结论. 同时对定理 3.2.12 的证明稍作修改，可直接得到关于下鞅(上鞅)同样的结论. ■

可选抽样定理在金融市场的具体含义为：若投资者持有某项金融工具的多头头寸，该工具的价格过程是一个鞅，即在已知到现在为止的所有市场信息条件下得到的未来价格的条件期望值等于它当前的市场价格，那么任一停止交易策略均不改变价格过程的鞅性质.

3.2.2.3 最优停时

设 X_t 为某个金融工具在 t 时刻的价值，$t=1$，…，T. 则在 T 时刻，可知其价值为 X_T，并可知道该投资是否盈利. 在 $T-1$ 时刻，比较 X_{T-1}和未来价格 X_T 关于$\mathcal{F}_{T-1}$的条件期望，如果 $X_{T-1} \geqslant E(X_T \mid \mathcal{F}_{T-1})$，人们的理性会选择结束该投资，反之则继续持有. 那么，这种退出交易的策略是否是最优的呢?

定义 3.2.19(Snell 包络) 设$\{X_t : t=1, \cdots, T\}$是关于滤子$\{\mathcal{F}_t : t=1, \cdots, T\}$的适应过程，且满足

$$E|X_t| < \infty, \quad t = 1, \cdots, T,$$

定义下列的递归过程：

$$Z_T = X_T,$$
$$Z_t = \max\{X_t, E(Z_{t+1} \mid \mathcal{F}_t)\}, \quad t = 0, \cdots, T-1,$$

称该过程为$\{X_t\}$的 **Snell 包络.**

由定义知：Snell 包络$\{Z_t\}$不小于$\{X_t\}$，即

95

$$Z_t \geqslant X_t, \quad t = 0, \cdots, T.$$

且有

$$Z_t \geqslant E(Z_{t+1} \mid \mathcal{F}_t), \quad t = 0, \cdots, T,$$

即$\{Z_t\}$是一个上鞅.

命题 3.2.20 Snell 包络$\{Z_t : t=1, \cdots, T\}$是不小于$\{X_t : t=1, \cdots, T\}$的最小上鞅.

证明 设$\{Y_t\}$是不小于$\{X_t\}$的任意一个上鞅，可由倒向归纳法证明

$$Y_t \geqslant Z_t, \quad t = 1, \cdots, T.$$

事实上，对 $t=T$，根据 Snell 包络的定义我们有 $Y_T = Z_T = X_T$，现在假设 $Y_t \geqslant Z_t$，则由上

鞅的性质有

$$Y_{t-1} \geqslant E(Y_t \mid \mathcal{F}_{t-1}) \geqslant E(Z_t \mid \mathcal{F}_{t-1}).$$

又因为$\{Y_t\}$不小于$\{X_t\}$，故 $Y_{t-1} \geqslant X_{t-1}$，从而

$$Y_{t-1} \geqslant \max\{X_{t-1}, E(Z_t \mid \mathcal{F}_{t-1})\} = Z_{t-1}.$$

■

定义停时

$$\tau^* = \min\{t \geqslant 0 : Z_t = X_t\},$$

它表示满足 $Z_t = X_t$ 的首达时刻. 注意到

$$\{\tau^* = k\} = \{X_0 < Z_0, \cdots, X_{k-1} < Z_{k-1}, X_k = Z_k\},$$

且由 $Z_T = X_T$ 知 τ^* 取值于$\{0, \cdots, T\}$.

定理 3.2.21 停时 Snell 包络 $Z_t^{\tau^*} = Z_{\min(t,\tau^*)}$ $(t=1, \cdots, T)$是一个$\mathcal{F}_t$ 鞅.

证明 只需证明 $E(Z_{t+1}^{\tau^*} - Z_t^{\tau^*} \mid \mathcal{F}_t) = 0$ $(t=0, \cdots, T-1)$即可. 事实上，根据引理 3.2.17，停时过程 $Z_t^{\tau^*}$ 可以写成下列形式的离散时间随机积分，对每个 t，均有

96

$$\begin{aligned} Z_t^{\tau^*} &= Z_{\min(\tau^*, t)} \\ &= Z_0 + \sum_{i=1}^{t} \mathbf{1}(i \leqslant \tau^*)(Z_{i+1} - Z_i). \end{aligned}$$

因此

$$Z_{t+1}^{\tau^*} - Z_t^{\tau^*} = \mathbf{1}(t+1 \leqslant \tau^*)(Z_{t+1} - Z_t).$$

对上式有下列结论：若用 $E(Z_{t+1} \mid \mathcal{F}_t)$替换 Z_t，上式仍然成立，即

$$Z_{t+1}^{\tau^*} - Z_t^{\tau^*} = \mathbf{1}(t+1 \leqslant \tau^*)(Z_{t+1} - E(Z_{t+1} \mid \mathcal{F}_t)) \tag{3.3}$$

对 $t=0, \cdots, T-1$ 成立. 实际上，对固定的 t，在集合$\{t+1 \leqslant \tau^*\}^c = \{\tau^* \leqslant t\}$上，上式右边为零. 又由 $Z_{t+1}^{\tau^*} = Z_{\min(\tau^*, t+1)} = Z_{\tau^*}$ 和 $Z_t^{\tau^*} = Z_{\tau^*}$ 知，上式左边在该集合上也为零. 而在补集$\{t+1 \leqslant \tau^*\} = \{X_1 < Z_1, \cdots, X_t < Z_t\}$上，有

$$Z_t = \max\{X_t, E(Z_{t+1} \mid \mathcal{F}_t)\} = E(Z_{t+1} \mid \mathcal{F}_t),$$

从而

$$\begin{aligned} Z_{t+1}^{\tau^*} - Z_t^{\tau^*} &= \mathbf{1}(t+1 \leqslant \tau^*)(Z_{t+1} - Z_t) \\ &= \mathbf{1}(t+1 \leqslant \tau^*)(Z_{t+1} - E(Z_{t+1} \mid \mathcal{F}_t)). \end{aligned}$$

综上所述，(3.3)成立.

因为 $\mathbf{1}(t+1 \leqslant \tau^*) = \mathbf{1}_{\{\tau^* \leqslant t\}^c}$ 是$\mathcal{F}_t$ 可测的，所以

$$\begin{aligned} E(Z_{t+1}^{\tau^*} - Z_t^{\tau^*} \mid \mathcal{F}_t) &= \mathbf{1}(t+1 \leqslant \tau^*) E((Z_{t+1} - E(Z_{t+1} \mid \mathcal{F}_t)) \mid \mathcal{F}_t) \\ &= \mathbf{1}(t+1 \leqslant \tau^*)(E(Z_{t+1} \mid \mathcal{F}_t) - E(Z_{t+1} \mathcal{F}_t)) \\ &= 0, \end{aligned}$$

定理得证. ■

现在考虑不大于 T 的所有停时中使 $E(X_\tau)$最大的停时. 记$\mathcal{T}_{k,T}$为所有取值于$\{k, \cdots, T\}$的停时的集合，由 $Z_T = X_T$ 知 $\tau^* \leqslant T$，于是 $\tau^* \in \mathcal{T}_{0,T}$

97

定义 3.2.22 称停时 $\tau \in \mathcal{T}_{0,T}$ 为$\{X_t : t=1, \cdots, T\}$的**最优停时**，如果它满足下列条件

$$E(X_\tau)=\sup\{E(X_\sigma):\sigma\in\mathcal{T}_{0,T}\}.$$

关于最优停时，其主要结果为下列定理.

定理 3.2.23 如 τ^* 是 $\{X_t:t=1,\cdots,T\}$ 的最优停时，则

$$Z_0=E(X_{\tau^*})=\sup\{E(X_\sigma):\sigma\in\mathcal{T}_{0,T}\}.$$

证明 因为停时过程 $\{Z_t^{\tau^*}:t=1,\cdots,T\}$ 是鞅，则 Z_0 不依赖于 $t=0,\cdots,T$，即满足

$$Z_0=E(Z_0^{\tau^*})=E(Z_T^{\tau^*})$$

又因为 $\tau^*\leqslant T$，所以

$$Z_T^{\tau^*}=Z_{\min(\tau^*,T)}=Z_{\tau^*}.$$

再由 τ^* 的定义有 $Z_{\tau^*}=X_{\tau^*}$. 所以

$$E(X_{\tau^*})=E(Z_{\tau^*})=E(Z_T^{\tau^*})=Z_0.$$

设 $\tau\in\mathcal{T}_{0,T}$ 是任意一个停时，如果 $E(X_\tau)\leqslant E(X_{\tau^*})$，那么 τ^* 是最优停时. 因为 Snell 包络 $\{Z_t:t=1,\cdots,T\}$ 是上鞅，所以停时过程 $\{Z_t^\tau\}$ 是上鞅. 于是有

$$Z_0=Z_0^\tau\geqslant E(Z_T^\tau)=E(Z_\tau),$$

最后一个等式是由 $\tau\leqslant T$，得到 $Z_{\min(\tau,T)}=Z_\tau$. 又因为 $\{Z_t\}$ 不小于 $\{X_t\}$，所以 $Z_\tau\geqslant X_\tau$（关于 ω 点态），则 $E(Z_\tau)\geqslant E(X_\tau)$，最后，由 τ^* 的最优性有

$$Z_0=\sup\{E(X_\sigma):\sigma\in\mathcal{T}_{0,T}\}.$$ ■

由定理 3.2.23 可以看出，τ^* 的确是最大化期望 $E(X_\tau)$ 的最优停时，而且，由上述证明过程可以看出，使用倒向归纳法能够计算出该最大值，因为它是 Snell 包络在 0 时刻的取值 Z_0.

3.2.2.4 极大不等式

本小节讨论一个极大不等式，它给出了到 n 时刻的部分和过程的绝对值超过某个固定数的概率的上界. 由切比雪夫不等式知

98

$$P(|S_n|\geqslant\varepsilon)=P\left(\left|\sum_{i=1}^n X_i\right|\geqslant\varepsilon\right)$$

不大于 $\mathrm{Var}(S_n)/\varepsilon^2=\sigma^2/(n\varepsilon^2)$，其中 $S_n=\sum_{i=1}^n X_i, X_1,\cdots,X_n$ 是独立同分布随机变量序列，$E(X_1)=0$，$\sigma^2=\mathrm{Var}(X_1)\in(0,\infty)$. 这意味着，一个随机游动（如价格过程）在 n 时刻偏离其均值超过 ε 的概率可以被其标的 X_i 的二阶矩有效控制. 不禁要问，该结论对到 n 时刻为止的最大偏移还成立吗？或 $P(\max_{1\leqslant k\leqslant n}S_k>\varepsilon)$ 是否仍有界？称这类不等式为极大不等式.

引理 3.2.24 设 $\{X_n:n\in\mathbf{N}_0\}$ 是一个下鞅，N 为停时，且有 $P(N\leqslant n)=1$，$n\in\mathbf{N}$，n 为某个固定数，则 $E(X_0)\leqslant E(X_N)\leqslant E(X_n)$.

证明 先证明第一个不等式，由 $P(N\leqslant n)=1$ 知

$$E|X_N|\leqslant E\left(\max_{1\leqslant k\leqslant n}|X_i|\right)\leqslant\sum_{i=1}^n E|X_i|<\infty.$$

又因为

$$X_{\min(N,n)} = X_n\mathbf{1}(N>n)+X_N\mathbf{1}(N\leqslant n),$$

所以

$$\begin{aligned} X_N &= X_N\mathbf{1}(N\leqslant n)+X_N\mathbf{1}(N>n) \\ &= (X_{\min(N,n)}-X_n\mathbf{1}(N>n))+X_N\mathbf{1}(N>n). \end{aligned}$$

因为 $X_{\min(N,n)}(n\in\mathbf{N})$是下鞅，故 $E(X_{\min(N,n)})\geqslant EX_0$，在上式两边同时取期望得

$$E(X_N)\geqslant E(X_0)-E(X_n\mathbf{1}(N>n))+E(X_N\mathbf{1}(N>n)).$$

令 $n\to\infty$，由控制收敛定理知

$$E(X_N)\geqslant E(X_0)+\lim_{n\to\infty}E(X_N\mathbf{1}(N>n))=E(X_0),$$

第一个不等式得证.

现证第二个不等式. 设 $\varphi_n=\mathbf{1}(N<n)=\mathbf{1}(N\leqslant n-1)$，$n\in\mathbf{N}$，$\varphi_0=0$. 则 φ_n 关于$\mathcal{F}_{n-1}$ 可测，且 [99]

$$\int_0^n\varphi_t\mathrm{d}X_t=\sum_{i=1}^n\mathbf{1}(N<i)(X_i-X_{i-1}).$$

当 $N\geqslant n$ 时，有 $\int_0^n\varphi_t\mathrm{d}X_t=0$. 当 $N<n$ 时，

$$\int_0^n\varphi_t\mathrm{d}X_t=\sum_{i=N+1}^n(X_i-X_{i-1})=X_n-X_N,$$

于是有

$$\int_0^n\varphi_t\mathrm{d}X_t=X_n-X_{\min(n,N)}.$$

因为上式左边的离散随机积分是下鞅，所以 $X_n-X_{\min(n,N)}$ 是一个下鞅. 从而得到下列的结论

$$0=E\left(\int_0^0\varphi_t\mathrm{d}X_t\right)\leqslant E\left(\int_0^n\varphi_t\mathrm{d}X_t\right)=\begin{cases}0, & n\geqslant N,\\ E(X_n)-E(X_N), & n<N.\end{cases}$$

所以 $E(X_N)\leqslant E(X_n)$. ■

定理 3.2.25(Doob 极大不等式)

设$\{X_n\}$是一个下鞅，从而有 $E(X_{n+1}\mid\mathcal{F}_n)\geqslant X_n$($P$-a. s.)，$c>0$ 为一个常数. 则 $X_n^+=\max(0, X_n)$满足不等式

$$cP\left(\max_{0\leqslant k\leqslant n}X_k^+\geqslant c\right)\leqslant E\left(X_n1_{\{\max_{0\leqslant k\leqslant n}X_k^+\geqslant c\}}\right)\leqslant EX_n^+.$$

特别地，有

$$P\left(0\max_{0\leqslant k\leqslant n}X_k^+\geqslant c\right)\leqslant\frac{E(X_n^+)}{c}.$$

证明 设 $N=\inf\{1\leqslant k\leqslant n: X_k\geqslant c\}$，规定 $\inf\varnothing=0$. 因此，N 是 X_k 到达 c 的最小时刻，如不能到达，则取值为 0. 在集合$\{N\leqslant n\}$和

$$A = \left\{ \max_{0\leqslant k\leqslant n} X_k^+ \geqslant c \right\} = \left\{ \max_{0\leqslant k\leqslant n} X_k \geqslant c \right\}$$

的交集上，有 $X_N \geqslant c$. 因此得到 $E(X_N 1_A) \geqslant E(c1_A) = c \cdot P(A)$. 最后只需证明 $E(X_N 1_A) \leqslant$ [100] $E(X_n^+) = E(\max(0, X_n))$即可. 事实上，由引理 3.2.24 有 $E(X_n) \geqslant E(X_N)$. 因此

$$E(X_n) \geqslant E(X_N 1_A) + E(X_N 1_{A^c}).$$

在 A^c 上有 $N = n$，从而 $E(X_N 1_{A^c}) = E(X_n 1_{A^c})$，这等价于 $E(X_n 1_A) \geqslant E(X_N 1_A)$，再由 X_n^+ 的定义知，$X_n 1_A \leqslant X_n^+ 1_A$，所以 $E(X_n^+) \geqslant E(X_N 1_A)$. ■

下面介绍 Kolmogorov 不等式，它是切比雪夫不等式的推广，同时也是 Doob 极大不等式的特例.

推论 3.2.26(Kolmogorov 不等式) 设 ξ_1, …, ξ_n 是 i.i.d.，且 $E(\xi_1) = 0$，$E(\xi_1^2) < \infty$. 设 $S_n = \sum_{i=1}^n \xi_i$，那么 S_n^2 是一个下鞅，且对任意的 $c = \varepsilon^2$，$\varepsilon > 0$，有

$$P\left(\max_{1\leqslant k\leqslant n} |S_k| \geqslant \varepsilon \right) = P\left(\max_{1\leqslant k\leqslant n} |S_k^2| \geqslant c \right) \leqslant \frac{E(S_n^2)}{c} = \frac{E(S_n^2)}{\varepsilon^2}.$$

3.2.3 推广到 $\mathbf{R}^d$ 值过程

本小节把随机过程$\{X_n\}$取值于 $\mathbf{R}$ 的定义和结论推广到$\{X_n\}$取值于 $\mathbf{R}^d$ 的情形，其中整数 $d > 1$. 首先定义 $\mathbf{R}^d$ 空间的范数 $\|\cdot\|$，设有一列随机变量 $X_n:(\Omega, \mathcal{F}, P) \to (\mathbf{R}^d, \mathcal{B}^d)$，其中 $\mathcal{B}^d$ 表示 $\mathbf{R}^d$ 上的 Borel σ 代数. 如果$\{X_n\}$是$\mathcal{F}_n$ 适应过程，且有 $E\|X_n\| < \infty$，$E(X_{n+1} | \mathcal{F}_n) = X_n$ 对所有 n a.s. 成立，则称$\{X_n\}$是$\mathcal{F}_n$ 鞅. 对任意一个随机向量 $X = (X_1, \cdots, X_d)'$和子 σ 代数$\mathcal{A}$，定义 $E(X|\mathcal{A}) = (E(X_1|\mathcal{A}), \cdots, E(X_d|\mathcal{A}))'$. 可预报过程的定义可作类似的推广.

定义 3.2.27 设$\{X_n\}$是$\mathcal{F}_n$ 适应过程，取值于 $\mathbf{R}^d$，$\{\varphi_n\}$是取值于 $\mathbf{R}^d$ 的可预报过程，称取值于 $\mathbf{R}$ 的过程 $I_n:(\Omega, \mathcal{F}, P) \to (\mathbf{R}, \mathcal{B})$，

$$I_n = \sum_{i=1}^n \varphi_i'(X_i - X_{i-1}), \quad n \geqslant 0$$

为**(多维)离散随机积分**，记为 $I_n = \int_0^n \varphi_t' \mathrm{d}X_t$.

将定理 3.2.12 推广到多维情形，有下列结论.

定理 3.2.28 设$\{X_n\}$是$\mathcal{F}_n$(下)鞅，$\{\varphi_n\}$是$\mathcal{F}_n$ 适应过程. 如果 $\|\varphi_n\| \in L_p$，$\|X_n\| \in L_q$，其中 p，$q \in [0, \infty]$且$\frac{1}{p} + \frac{1}{q} = 1$，则 I_n 是$\mathcal{F}_n$(下)鞅，且对任意一个常数 X_0，$X_0 + I_n$ [101] 也是一个(下)鞅.

证明 由 Hölder 不等式，有

$$E\|\varphi_i' X_i\| \leqslant E(\|\varphi_i\| \cdot \|X_i\|) \leqslant \left(\int \|\varphi_i\|^p \mathrm{d}P \right)^{1/p} \cdot \left(\int \|X_i\|^q \mathrm{d}P \right)^{1/q} < \infty.$$

因此，

$$E|X_0 + I_n| \leqslant E|X_0| + \sum_{i=1}^n E\|\varphi_i' X_i\| + \|\varphi_i' X_{i-1}\| < \infty,$$

余下证明类似于定理 3.2.12 的方法. ■

3.3 平稳序列

对一个时间序列 X_1, X_2, …, 有时需将其观测值与其他观测值区分, 比如, 在满足哪些最少的假设下, 样本均值的计算公式有意义?

设前 T 个观测值为 X_1, …, X_T, 如果它们服从相同的分布, 那么对 $t=1$, …, T, 均有 $E(X_t)=E(X_1)$, 此时样本均值是有意义的, 可以用它来对随机波动和总体均值进行有效的估计.

对 $E(X_t)=0$, $t=1$, …, T, 有时需要估计某些至少与 2 个变量相关的统计量, 比如, 连续观测值的协方差 $E(X_tX_{t+1})$. 要能用求平均值方法来估计此期望, 交叉矩必须不随时间 $t=1$, …, $T-1$ 变化, 即连续观测值对

$$(X_1, X_2), (X_2, X_3), \cdots \tag{3.4}$$

有相同的交叉乘积期望

$$E(X_1X_2) = E(X_2X_3) = \cdots.$$

满足此性质的一个充分条件是: (3.4)中的观测值对服从相同的分布. 这就引出下列严平稳和弱平稳的概念, 它们是刻画离散时间序列关系的两个重要概念.

3.3.1 弱平稳和严平稳

定义 3.3.1(弱平稳) 设 $\{X_t: t\in \mathbf{Z}\}$ 是取值于 $\mathbf{R}$ 或 $\mathbf{C}$ 的时间序列.

(ⅰ) 称 $\{X_t\}$ 是**严平稳的**, 如果对所有的 t_1, …, $t_n\in\mathbf{Z}$ 和 $h\in\mathbf{Z}$, 均有

$$(X_{t_1}, \cdots, X_{t_n}) \stackrel{d}{=} (X_{t_1+h}, \cdots, X_{t_n+h}).$$

(ⅱ) 称 $\{X_t\}$ 为 **L_2 时间序列**, 如果

$$E|X_t|^2 = E(X_t\overline{X}_t) < \infty, \quad t\in\mathbf{Z}.$$

102

(ⅲ) 称一个 L_2 序列是**(弱)平稳的**, 如果

$$E(X_t) = E(X_1), \quad t\in\mathbf{Z},$$

且对所有的 $h\in\mathbf{Z}$, 有

$$E(X_tX_{t+h}) = E(X_1X_{1+h}).$$

注 3.3.2

(ⅰ) 术语"平稳"一般指的是弱平稳序列.

(ⅱ) 严平稳意味着序列关于**时间推移**(或**滞后**)具有相同的有限维分布. 设 $h\in\mathbf{Z}$, 考虑下列的 h 推移算子

$$L_h(\{X_t\}) = \{X_{t+h}: t\in\mathbf{Z}\},$$

同时用

$$\Pi_{t_1,\cdots,t_n}(\{X_t\}) = (X_{t_1}, \cdots, X_{t_n})$$

表示序列在 t_1, …, t_n 上的投影, $\{X_t\}$ 是严平稳的, 当且仅当

$$\Pi_{t_1,\cdots,t_n}(L_h(\{X_t\}))\stackrel{d}{=}\Pi_{t_1,\cdots,t_n}(\{X_t\})$$

对所有的 t_1，…，$t_n\in\mathbf{Z}$ 和 $h\in\mathbf{Z}$ 成立.

（iii）尽管很多结论仅需要弱平稳的条件就可以了，但是在实际应用中很多问题的解决至少有一步要用到中心极限定理，这就要求比弱平稳更严格的条件，如严平稳或要求序列是由具有独立同分布的误差项的序列导出的时间序列.

定义 3.3.3 设 $\{X_t\}$ 是一个弱平稳的时间序列，则函数

$$\gamma_X(h)=\operatorname{Cov}(X_{t+h},X_t),\quad h\in\mathbf{Z}$$

有定义，称为**自协方差函数**. 若 $\gamma_X(0)>0$，称相关函数

$$\rho_X(h)=\operatorname{Cor}(X_1,X_{1+h})=\frac{\gamma_X(h)}{\gamma_X(0)},\quad h\in\mathbf{Z}$$

为**自相关函数(ACF)**.

103

注 3.3.4 记 L_2 中的随机子向量 $(X_1,\cdots,X_T)$ 的 $T\times T$ 协方差矩阵为

$$\Gamma_X=(\operatorname{Cov}(X_i,X_j))_{i,j},$$

如果矩阵只与维数 T 有关，则记为 Γ_T. 对一个弱平稳序列，其协方差矩阵为

$$\Gamma_T=\begin{pmatrix}\gamma(0) & \gamma(1) & \cdots & \gamma(T-1)\\ \gamma(1) & \gamma(0) & \cdots & \gamma(T-2)\\ \vdots & \ddots & \ddots & \vdots\\ \gamma(T-1) & \cdots & \gamma(1) & \gamma(0)\end{pmatrix}.$$

下面是若干个例子.

▶**例 3.3.5**

（i）设 $\varepsilon_t(t\in\mathbf{Z})$ 是一列 i.i.d. 随机变量，且期望为 0，方差为 $\sigma^2\in(0,\infty)$，则 $\{\varepsilon_t\}$ 是弱平稳的. 定义一阶差分序列

$$X_t=\varepsilon_t-\varepsilon_{t-1},\quad t\in\mathbf{Z},$$

则 $E(X_t)=0$，若 $i=j$，$E(\varepsilon_i\varepsilon_j)=\sigma^2$，否则为 0. 于是有

$$\begin{aligned}E(X_tX_{t+h})&=E[(\varepsilon_t-\varepsilon_{t-1})(\varepsilon_{t+h}-\varepsilon_{t+h-1})]\\&=E(\varepsilon_t\varepsilon_{t+h})-E(\varepsilon_{t-1}\varepsilon_{t+h})-E(\varepsilon_t\varepsilon_{t+h-1})+E(\varepsilon_{t-1}\varepsilon_{t+h-1})\\&=\begin{cases}2\sigma^2, & h=0,\\ -\sigma^2, & h=-1,1,\\ 0, & |h|>1.\end{cases}\end{aligned}$$

因此，$\{X_t\}$ 也是弱平稳的. 其相关 ACF 为

$$\rho_X(h)=\begin{cases}1, & h=0,\\ -1/2, & |h|=1,\\ 0, & |h|>1.\end{cases}$$

（ii）设 $\{X_t:t\in\mathbf{Z}\}$ 是一个弱平稳的时间序列，自协方差函数为

$$\gamma(h)=E(X_0X_h),\quad h\in\mathbf{Z}.$$

记

$$Y_t := X_t + X_0, t \in \mathbf{Z}.$$

易证

$$\operatorname{Cov}(Y_s, Y_t) = \gamma(0) + \gamma(s) + \gamma(t) + \gamma(s-t).$$

所以$\{Y_t\}$不是弱平稳的.

104

(ⅲ) 下列例子看起来有点出人意料. 设 X 是一个随机变量, $X \stackrel{d}{=} -X$, $\sigma^2 = \operatorname{Var}(X) \in (0, \infty)$, 定义序列

$$X_t := (-1)^t X, \quad t \in \mathbf{N}.$$

则$\{X_t : t \geqslant 1\}$是弱平稳的. 实际上, 由 $E(X)=0$ 知

$$E(X_t) = 0, \quad t \in \mathbf{N},$$

且有

$$\begin{aligned}\operatorname{Cov}(X_t, X_{t+k}) &= \operatorname{Cov}((-1)^t X, (-1)^{t+k} X) \\ &= (-1)^{2t+k} \sigma^2 \\ &= \begin{cases} \sigma^2, & k \text{ 是偶数}, \\ -\sigma^2, & k \text{ 是奇数}, \end{cases}\end{aligned}$$

因此, $\gamma_X(t, t+k)$不依赖于 t.

(ⅳ) 非平稳时间序列可能存在平稳的子列. 例如, 考虑下列周期性序列:

$$X_t = \sin\left(\frac{2\pi}{r} t\right) + \varepsilon_t, \quad t \in \mathbf{Z},$$

其中$\{\varepsilon_t\}_{t \in \mathbf{Z}}$是弱(严)平稳的, 趋势函数项有周期 r, 对 $0 \leqslant h < r$, 定义

$$Y_t = X_{t \cdot r + h}, \quad t \in \mathbf{Z},$$

则$\{Y_t\}_{t \in \mathbf{Z}}$是弱(严)平稳的.

(ⅴ) 下列例子表明, 可以由均值和方差来判断一个时间序列的平稳性, 但它并不是一个充分的条件. 设 X_1, X_2, …是 i. i. d., $E(X_1)=0$ 且 $\sigma^2 = \operatorname{Var}(X_1) \in (0, \infty)$. 考虑随机游动

$$S_n = \sum_{i=1}^{n} X_i, \quad n = 1, 2, \cdots.$$

显然有 $E(S_n)=0$ 且 $\operatorname{Var}(S_n) = n\sigma^2$, 所以$\{S_n : n \in \mathbf{N}\}$不是弱平稳的, 因为它的方差与参数 n 有关. 现在考虑变量

$$S_n^* = \frac{S_n}{\sqrt{n}}, \quad n = 1, 2, \cdots.$$

105

显然 $E(S_n^*)=0$ 且 $\operatorname{Var}(S_n^*) = \sigma^2$. 下面来观察 S_n^* 是否是平稳的, 因为

$$\begin{aligned}\operatorname{Cov}(S_n^*, S_m^*) &= \frac{1}{\sqrt{nm}} E(S_n S_m) \\ &= \frac{1}{\sqrt{nm}} E\left[\left(\sum_{i=1}^{n} X_i\right)\left(\sum_{i=1}^{m} X_i\right)\right] \\ &= \frac{1}{\sqrt{nm}} E\left(\sum_{i=1}^{n} \sum_{j=1}^{m} X_i X_j\right)\end{aligned}$$

$$= \frac{1}{\sqrt{nm}} \sum_{i=1}^{n} \sum_{j=1}^{m} E(X_i X_j).$$

因此，当 $i=j$ 时有 $E(X_iX_j)=\sigma^2$，否则 $E(X_iX_j)=0$. 则

$$\operatorname{Cov}(S_n^*, S_m^*) = \frac{\sigma^2}{\sqrt{nm}} \min(n,m),$$

其中 n，$m\in\mathbf{N}$，上式不是 $n-m$ 的函数，因此$\{S_n^*\}$不是弱平稳的. ◀

引理 3.3.6 严平稳 L_2 时间序列是弱平稳的，但存在不是严平稳的弱平稳序列.

证明 由严平稳序列$\{X_t\}$的定义知

$$X_1 \overset{d}{=} X_t \quad 和 \quad (X_t, X_{t+h}) \overset{d}{=} (X_1, X_{1+h})$$

对所有的 t 成立. 由于 $E(X_1^2)<\infty$ 且 $\mu E|X_1|<\infty$，则有 $E(X_t)=E(X_1)$，且 $E(X_tX_{t+h})=E(X_1X_{1+h})$，$\forall t$，$h$. 因此它是弱平稳的. 下面给出弱平稳不是严平稳的一个反例. 设独立随机变量 X_t 具有下列边际分布

$$若\ t\ 是偶数，\quad 则\ X_t \sim N(0,1/3),$$
$$若\ t\ 是奇数，\quad 则\ X_t \sim U(-1,1),$$

显然，$\{X_t\}$不是严平稳的. 但是对所有 t 有 $E(X_t)=0$，$\operatorname{Var}(X_t)=1/3$，对 $h>0$ 有 $\gamma_X(h)=0$，从而$\{X_t\}$是弱平稳的. ■

引理 3.3.7 弱平稳过程的自协方差函数满足下列性质：

（ⅰ）$\gamma(0)$是非负实数.

106

（ⅱ）$\gamma(-h)=\overline{\gamma(h)}$对 $h\in\mathbf{Z}$ 成立.

（ⅲ）$\forall h\in\mathbf{Z}$，有$|\gamma(h)|\leqslant\gamma(0)$.

（ⅳ）若序列是一个不相关的时间序列，则当$|h|\geqslant 1$ 时，有 $\gamma(h)=0$.

证明 设$\{X_t\}$是弱平稳序列.

（ⅰ）显然有 $\gamma(0)=E((X_1-\mu)\overline{(X_1-\mu)})=E|X_1|^2\in[0,\infty)$.

（ⅱ）设 $\mu=E(X_1)$，则对所有的 h 有

$$\begin{aligned}
\gamma(-h) &= E((X_{t-h}-\mu)(\overline{X_t-\mu})) \\
&= E(X_t-\mu)(\overline{X_{t+h}-\mu}) \\
&= E\left(\overline{(\overline{X_t-\mu})(X_{t+h}-\mu)}\right) \\
&= \overline{E(\overline{X_t-\mu})(X_{t+h}-\mu)} \\
&= \overline{\gamma(h)}.
\end{aligned}$$

（ⅲ）由 Cauchy-Schwarz 不等式有

$$\begin{aligned}
|\gamma(h)| &= |\operatorname{Cov}(X_1, X_{1+h})| \\
&\leqslant \sqrt{\operatorname{Var}(X_1)\operatorname{Var}(X_{1+h})} \\
&= \sqrt{\gamma(0)^2} \\
&= \gamma(0).
\end{aligned}$$

(ⅳ) 显然，对 $h\neq 0$ 有 $\gamma(h)=\mathrm{Cov}(X_1, X_{1+h})=0$ ■

定义 3.3.8(白噪声)　设$\{\varepsilon_t\}$是两两互不相关的随机变量序列，且当 $t\neq s$ 时，有

$$\mathrm{Cov}(\varepsilon_t,\varepsilon_s)=0,\quad E(\varepsilon_t)=0,\quad \mathrm{Var}(\varepsilon_t)=\sigma^2\in(0,\infty),$$

则称$\{\varepsilon_t\}$为**白噪声过程**(或**序列**)，简称**白噪声**，记为 $\varepsilon_t\sim \mathrm{WN}(0, \sigma^2)$.

注 3.3.9

(ⅰ) 白噪声过程是弱平稳的，且自协方差函数为

$$\gamma(h)=\begin{cases}\sigma^2, & h=0,\\ 0, & \text{其他情形}.\end{cases}$$

[107]

(ⅱ) $\varepsilon_t\sim \mathrm{WN}(0, \sigma^2)$只指定自协方差函数，没有指定具体的边际分布和多维联合分布，也没有要求 ε_t 是相互独立的，所以 $\varepsilon_t\sim \mathrm{WN}(0, \sigma^2)$的条件比 $\varepsilon_t\overset{\text{i.i.d.}}{\sim}N(0, \sigma^2)$更弱，更一般.

(ⅲ) 由定义知，$\mathcal{F}_t$ 鞅差序列$\{\varepsilon_t\}$满足 $E|\varepsilon_t|<\infty$ 和 $E(\varepsilon_t|\mathcal{F}_{t-1})=0$. 如果 ε_t 是 L_2 序列，则由引理 3.2.9，$\{\varepsilon_t\}$是不相关的. 因此，一个 L_2MDS 是一个白噪声过程，但是一个 MDS 有比白噪声过程更强的性质，例如条件期望 $E(\varepsilon_t|\mathcal{F}_{t-1})$等于 0，当然也有 $E(\varepsilon_t)=0$. 而白噪声过程仅有 $E(\varepsilon_t)=0$，没有性质 $E(\varepsilon_t|\mathcal{F}_{t-1})=0$.

设$\{X_t\}$是一个弱平稳过程，记 $\mu=E(X_1)$，$\{X_t\}$的自协方差函数在讨论序列的相关性时起着重要的作用. 由弱平稳性知，随机变量

$$(X_t-\mu)(X_{t+h}-\mu),\quad t=1,\cdots,T-h$$

有相同的期望，记其为 $\gamma(h)$. 考虑取平均得到的估计量

$$\widetilde{\gamma}_T(h)=\frac{1}{T-h}\sum_{t=1}^{T-h}(X_{t+h}-\mu)(X_t-\mu),$$

其中 μ 为上面的期望，且是已知的. 直接验证得到：$E(\widetilde{\gamma}_T(h))=\gamma(h)$. 因此，无论 $\gamma(h)$ 为何值，$\widetilde{\gamma}_T(h)$都是 $\gamma(h)$的无偏估计量. 因为 μ 在实际中是未知的，通常用样本均值来代替.

定义 3.3.10　设$\{X_t\}$是一个时间序列，则称

$$\widetilde{\gamma}_T(h)=\frac{1}{T}\sum_{t=1}^{T-h}(X_{t+h}-\overline{X}_T)(X_t-\overline{X}_T),\quad 0\leqslant h\leqslant T-1,$$

$$\hat{\gamma}_T(h)=0,\quad T-1<h,$$

[108]

为**样本自协方差函数**. 如$\hat{\gamma}_T(0)>0$，则称

$$\hat{\rho}_T(h):=\frac{\hat{\gamma}(h)}{\hat{\gamma}(0)},\quad h\geqslant 0$$

为**样本自相关函数**(ACF).

注意到样本自协方差函数的定义中是 $T-h$ 项求和，但是分母是 $1/T$ 而非 $1/(T-h)$. 这在时间序列分析中是很常见的，通常 T 比较大，所以差别可以忽略.

现在可以得到自协方差矩阵 Γ_T 的估计式为

$$\hat{\Gamma}_T = \begin{pmatrix} \hat{\gamma}_T(0) & \hat{\gamma}_T(1) & \cdots & \hat{\gamma}_T(T-1) \\ \hat{\gamma}_T(1) & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \hat{\gamma}_T(1) \\ \hat{\gamma}_T(T-1) & \cdots & \hat{\gamma}_T(1) & \hat{\gamma}_T(0) \end{pmatrix}.$$

图 3-2 描绘了瑞士信贷对数收益的样本自相关函数，从图中可看出收益之间不存在显著相关关系，这符合资产当日收益的经典结论. 水平线表示样本自相关的置信区间，它们是建立在当 $T\to\infty$ 时有

109

$$\sqrt{T}[\hat{\gamma}_T(h)-\gamma(h)] \xrightarrow{d} N(0,\sigma^4),$$

$$\sqrt{T}[\hat{\rho}_T(h)-\rho(h)] \xrightarrow{d} N(0,1),$$

基础上的，其中要求 X_t 是 i. i. d.，其方差 $\sigma^2\in(0,\infty)$，且存在 $\delta>0$，使 X_t 的 $4+\delta$ 阶矩有限，见定理 8.8.3. 因此，检验 $H_0:\gamma(h)=0$ 和 $H_1:\gamma(h)\neq 0$ 的渐近显著性检验中

$$\varphi = \begin{cases} 1, & \text{若 } |\hat{\rho}(h)| > \Phi^{-1}(1-\alpha/2)/\sqrt{T}, \\ 0, & \text{其他情形}. \end{cases}$$

进一步地，$\rho(h)$ 的置信区间为

$$\hat{\rho}_T(h) \pm \frac{\Phi^{-1}(1-\alpha/2)}{\sqrt{T}}.$$

即对 $\alpha=0.05$，样本自相关函数 $\hat{\rho}_T(h)$ 介于 $-1.96/\sqrt{T}$ 和 $1.96/\sqrt{T}$ 之间.

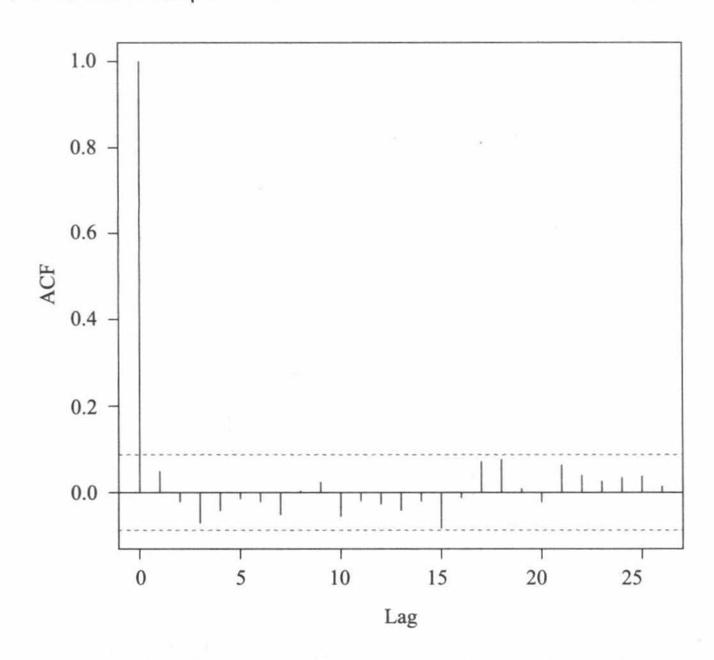


图 3-2 瑞士信贷对数收益的样本自相关函数(从 2/10/2003 到 9/1/2005)

引理 3.3.11 当 $T\geqslant 2$ 时，$\hat{\Gamma}_T$ 是半正定的.

下一步，考虑相关测度

$$P(X_t>0,X_{t+h}>0)=E(\mathbf{1}(X_t>0,X_{t+h}>0)). \tag{3.5}$$

如果 X_t 为 t 时刻的资产收益，该测度考虑相隔 h 的两期收益同取正值的联合概率，特别地，收益互相独立且分布关于零对称情形的概率为 1/4. 因为方程(3.5)不能写成均值和自协方差函数的函数，所以弱平稳不能保证方程(3.5)关于时滞 h 不变. 但是可以将方程(3.5)看成定义在集合

$$\mathcal{P}=\{\mathbf{P}:\mathbf{P}\text{ 是定义在 }\mathbf{R}^2\text{ 上的概率测度}\}$$

上的一个函数 τ，它是求 $P_{(X_t,X_{t+h})}\in\mathcal{P}$对应的函数值，即 $\tau:\mathcal{P}\to\mathbf{R}$ 的定义为：对 $\mathbf{P}\in\mathcal{P}$，

$$\tau(\mathbf{P})=\mathbf{P}(\{(x,y)\in\mathbf{R}^2:x>0,y>0\}).$$

如果 X_1，X_2，…是严平稳的，那么随机向量(X_1, X_{1+h})，(X_2, X_{2+h})，…具有相同的分布，因此，其中任意一个的分布都可代表其他随机向量的分布. 所以考虑下列估计量

$$\tilde{\tau}_T=\frac{1}{T-h}\sum_{t=1}^{T-h}\mathbf{1}(X_t>0,X_{t+h}>0)$$

是有意义的，该估计量显然是 $\tau(P_{(X_1,X_{1+h})})$的无偏估计量.

110

不禁要问：由严平稳序列导出的过程，例如 $\mathbf{1}(X_t>0, X_{t+h}>0)$，是否仍然是严平稳的？事实上，答案是肯定的.

命题 3.3.12 设 X_0，X_1，…是一个严平稳序列，且

$$g:\mathbf{R}^{\mathbf{N}_0}\to\mathbf{R},\quad (x_0,x_1,\cdots)\mapsto g(x_0,x_1,\cdots)$$

是一个可测映射. 那么序列 $Y_k=g(X_k, X_{k+1}, \cdots)(k\in\mathbf{N})$也是严平稳的.

证明 只需证明对所有可测集 $B\subset\mathbf{R}^{\mathbf{N}_0}$ 均有

$$P((Y_0,Y_1,\cdots)\in B)=P((Y_k,Y_{k+1,\dots})\in B).$$

设 $x=(x_0, x_1, \cdots)\in\mathbf{R}^{\mathbf{N}_0}$，对 $k\geqslant 1$，定义函数 $g_k(x)=g(x_k, x_{k+1}, \cdots)$. 对可测集 $B\subset\mathbf{R}^{\mathbf{N}_0}$，设

$$A=\{x:(g_0(x),g_1(x),\cdots)\in B\}.$$

由 A 的定义，有

$$\begin{aligned}P((Y_0,Y_1,\cdots)\in B)&=P((g_0(X_0,X_1,\cdots),g_1(X_0,X_1,\cdots),\cdots)\in B)\\&=P((X_0,X_1,\cdots)\in A)\\&=P((X_k,X_{k+1},\cdots)\in A)\\&=P((Yk,Y_{k+1},\cdots)\in B).\end{aligned}$$

■

3.4 线性过程和 ARMA 模型

经典时间序列分析的主要成果就是发展了 ARMA 模型，它是线性过程的特殊情形. 因此，本节从介绍线性过程开始，并介绍与线性过程相关的一些重要概念，如滞后算子和反演等，最后对 ARMA 模型作了详细的介绍.

3.4.1 线性过程和滞后算子

定义 3.4.1 设$\{X_n: n\in I\}$是一个普通概率空间上的随机变量集，参数集 $I=\mathbf{N}$ 或 $I=\mathbf{Z}$，$\{a_n: n\in I\}$为一个实数集. 称序列

111

$$Y_t = \sum_{i\in I} a_i X_{t-i}, \quad t \in I,$$

为**线性过程**或**线性滤波**. 如果当 $i<0$ 时，有 $a_i=0$，那么称对应的线性过程为**因果过程**.

注意，上述定义对序列$\{X_n\}$没有特别的要求，但在实际中通常考虑的是 i.i.d. 或者白噪声序列.

线性过程一个最简单的例子是下述时间序列$\{Y_t\}$：

$$Y_t = a_0 X_t + a_1 X_{t-1} + \cdots + a_q X_{t-q}, \quad t \in \mathbf{N},$$

它是 X_t，…，X_{t-q}的线性组合，称这种过程为 q 阶滑动平均过程，记为 MA(q).

线性滤波也可以通过**滞后算子**来表示，滞后算子在实数序列和随机变量序列中都有定义，其数学表达式为 $L(\{X_n\})=\{X_{n-1}\}$，其中$\{X_{n-1}\}=\{X_{n-1}: n\in\mathbf{Z}\}$表示滞后序列，即对任意 n，$L(\{X_n\})=\{Z_n\}$成立，当且仅当 $Z_n=X_{n-1}$. 通常把滞后序列的一般项元素也记成 $L(\{X_n\})$，尽管它有时是表示整个序列.

$$L(X_n) = LX_n = X_{n-1}, \quad n \in \mathbf{Z}.$$

显然，滞后算子是线性的，即它满足下列运算法则：

(ⅰ) 对所有的序列$\{X_n\}$和$\{Y_n\}$，有

$$L(X_n + Y_n) = L(X_n) + L(Y_n).$$

(ⅱ) 对任意的标量 a，有 $L(aX_n)=aL(X_n)$.

现考虑复合算子 L^2，

$$L^2(X_n) = L(L(X_n)) = L(X_{n-1}) = X_{n-2},$$

类似地，L 的任意 k 阶幂算子

$$L^k X_n = L(L^{k-1}(X_n)) = \cdots = X_{n-k},$$

其中 $k\geqslant 2$，规定 $L^0=1$(即恒等算子，满足 $1X_n=X_n$). 最后，可以引入负数幂

$$L^{-k} X_n = X_{n+k},$$

其中 $k\geqslant 1$. 因此，我们可以定义 L 的多项式，例如 $4L^3-L^2+2L+3$，它作用于$\{X_n\}$，得到

112

$$(4L^3 - L^2 + 2L + 3)(X_n) = 4X_{n-3} - X_{n-2} + 2X_{n-1} + 3X_n.$$

L 的多项式有明确的含义，称之为**滞后多项式**. 还可以定义**滞后序列算子**：

$$\theta(L) = \sum_{i\in\mathbf{N}} \theta_i L^i,$$

其中系数 $\theta_i\in\mathbf{C}$，该算子作用于$\{X_n\}$得到的结果为

$$\theta(L)X_n = (\theta(L)(\{X_n\}))_n = \sum_i \theta_i X_{n-i}.$$

当一个滞后序列由系数$\{\theta_i\}$表示时，作用在序列上的滞后多项式就是指 $\theta(L)$，记号 $\theta(L)$也通常指系数是$\{\theta_i\}$. 后面如无特殊说明，均默认这一约定.

已经看到，线性过程可以很方便地通过滞后多项式 $\theta(L)$ 写成 $Y_n=\theta(L)X_n$ 的形式，下面讨论，该序列是否是 L_2 收敛或者 a.s. 收敛的；如果 $\{X_n\}$ 是平稳的，$\{Y_n\}$ 是否也是平稳的？

考虑下列重要的因果过程，它是带白噪声扰动项序列的一个线性过程，

$$X_t=\sum_{k=0}^{\infty}\psi_k\varepsilon_{t-k},\quad \varepsilon_t\sim \mathrm{WN}(0,\sigma^2),$$

其中 $\sigma^2\in(0,\infty)$. 可证：如果条件 $\sum_{k=0}^{\infty}\psi_k^2<\infty$ 成立，那么 X_t 是平稳的. 事实上，如果条件成立，取截断序列 $X_t^{(n)}=\sum_{k=0}^{n}\psi_k\varepsilon_{t-k}$，当 $n\leqslant m$ 时，有

$$\begin{aligned}E(X_t^{(n)}-X_t^{(m)})^2&=\sum_{k,l=n+1}^{m}\psi_k\psi_l\mathrm{Cov}(\varepsilon_{t-k},\varepsilon_{t-l})\\&=\sigma^2\sum_{k=n+1}^{m}\psi_k^2,\end{aligned}$$

当 n，$m\to\infty$时，上式趋于零. 因此，$\{X_t^{(n)}:n\geqslant 1\}$是一个 L_2 柯西列，它的 L_2 极限为

$$X_t=\psi(L)\varepsilon_t=\sum_{k=0}^{\infty}\psi_k\varepsilon_{t-k}.$$

由白噪声的性质知 $E(X_t)=0$，对任意滞后长度 $h\in\mathbf{Z}$，

$$\begin{aligned}\gamma_X(h)&=\mathrm{Cov}(X_t,X_{t+h})\\&=\sum_{k,l=0}^{\infty}\psi_k\psi_l\mathrm{Cov}(\varepsilon_{t-k},\varepsilon_{t+h-l})\\&=\sigma^2\sum_{k=0}^{\infty}\psi_k\psi_{k+h}\end{aligned}$$

113

成立. 这表明 $X_t=\psi(L)\varepsilon_t$ 是平稳的. 值得一提的是，上述的收敛性对系数 ψ_k 是非增的情形显然成立，因为 $\psi_{k+h}\leqslant\psi_k$.

若$\{\varepsilon_t\}$是一个均值为 μ、自协方差为 $\gamma_\varepsilon(h)$的平稳时间序列，则 X_t 是平稳的，其均值为 $E(X_t)=\mu\sum_{k=0}^{\infty}\psi_k$，自协方差为

$$\gamma_X(h)=\sum_{k,l=0}^{\infty}\psi_k\psi_l\gamma_\varepsilon(l-k+h),\quad h\in\mathbf{Z}.$$

定义 3.4.2 如果 $\sum_i|\psi_i|<\infty$ 成立，则称滞后算子 $\psi(L)=\sum_i\psi_iL^i$ 为 L_1 **滤波**或 L_1 **绝对可求和**(可加的).

根据定理 A.5.1，对于 L_1 滤波有下列结论.

命题 3.4.3 设 $\psi(L)$是一个绝对可求和的线性滤波，$\{X_n\}$是随机变量序列，则

(ⅰ) 如果 $\sup_n E|X_n|<\infty$，那么 $Y_n=\psi(L)X_n$ 存在(a.s.)，且有 $E(Y_n)=\psi(L)\mu_n$，其中

$$\mu_n=E(X_n);$$

(ⅱ) 如果 $\sup_n EX_n^2<\infty$，那么 $Y_n=\psi(L)X_n$ 是 L_2 收敛的.

由上述结论知，如果 $\sum_i |\psi_i| < \infty$，那么可以用线性滤波 $\psi(L) = \sum_i \psi_i L^i$ 作用于一个白噪声过程$\{\varepsilon_t\}$，即 $X_t = \psi(L)\varepsilon_t$ 是(a. s.)存在的.

若有一个滞后算子 $\psi(L) = \sum_i \theta_i L^i$，则可以将其中的算子 L 替换成一个复变量 z.

定义 3.4.4 如果 $Y_n = \sum_i \theta_i X_{n-i}$ 是一个线性过程，其系数为$\{\theta_i\}$，对应的滞后序列为 $\psi(L) = \sum_i \theta_i L^i$，那么序列 $\psi(z) = \sum_i \theta_i z^i$，$z \in \mathbf{C}$ 称为一个 **z 变换**. 如果 $\psi(L)$具有下列形式：

$$\psi(L) = 1 - \theta_1 L - \cdots - \theta_p L^p,$$

其中 $p \in \mathbf{N}$，那么称其对应的 z 变换

$$\psi(z) = 1 - \theta_1 z - \cdots - \theta_p z^p, \quad z \in \mathbf{C}$$

114 为滤波的**特征函数**.

▶**例 3.4.5** 现介绍几个例子.

(i) 如果 $\psi(L) = \sum_{i=0}^{\infty} \theta_i L^i$ 是一个 L_1 滤波，即有 $\sum_{i=0}^{\infty} |\theta_i| < \infty$，那么幂级数 $\psi(z)$ 至少在单位圆 $E = \{z \in \mathbf{C}: |z| \leqslant 1\}$ 内收敛.

(ii) 如果 $p(L) = \sum_{i=-\infty}^{0} a_i L^i$，那么 $p(z) = a_0 + \frac{a_1}{z} + \frac{a_2}{z^2} + \cdots$ 是一个洛朗级数. 若$\{a_i\}$是绝对可求和的，则对所有满足 $\left|\frac{1}{z}\right| \leqslant 1 \Leftrightarrow |z| \geqslant 1$ 的 $z \in \mathbf{C}$，$p(z)$均存在. ◀

设有两个线性过程 $Y_n = \sum_i a_i X_{n-i}$ 和 $X_n = \sum_i b_i Z_{n-i}$，记

$$p(z) = \sum_i a_i z^i, \quad q(z) = \sum_i b_i z^i$$

分别为对应的 z 变换，则 X_n 与 Y_n 可以写成下列多项式形式

$$Y_n = p(L) X_n, \quad X_n = q(L) Z_n.$$

在 Y_n 的表达式中用 $q(L)Z_n$ 代替 X_n，得到

$$Y_n = p(L) X_n = p(L) q(L) Z_n.$$

即 $Y_n = r(L) Z_n$，其中 $r(L) = p(L) q(L)$，等价地有

$$r(z) = p(z) q(z),$$

其中 $p(z)$，$q(z)$为前面所给，可证该运算法则是正确的.

命题 3.4.6 设序列 $Y_n = p(L) X_n$，$X_n = q(L) Z_n$，其中 $p(L)$和 $q(L)$是滞后算子，它们的 z 变换在 $z \in \mathbf{C}$，$|z| \leqslant \rho (\rho > 0)$时收敛，则有 $Y_n = r(L) Z_n$，其中 $r(L)$的 z 变换为：

115
$$r(z) = p(z) q(z) = \sum_{i,j=0}^{\infty} a_i b_j z^{i+j}, \quad z \in \mathbf{C} \text{ 且 } |z| \leqslant \rho.$$

这意味着，复合算子 $p(L) \circ q(L)$具有下列形式

$$p(L) \cdot q(L) = \sum_{i,j=0}^{\infty} a_i b_j L^{i+j}.$$

证明 因为

$$
\begin{aligned}
Y_n &= \sum_i a_i X_{n-i} \\
&= \sum_{i=0}^{\infty} a_i \sum_{j=0}^{\infty} b_j Z_{n-(i+j)} \\
&= \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} a_i b_j L^{i+j} Z_n = \Big(\sum_{i,j=0}^{\infty} a_i b_j L^i L^j \Big) Z_n,
\end{aligned}
$$

这表明 $Y_n = r(L) Z_n$，其中 $r(L) = p(L) q(L)$. 根据定义，$r(L)$的 z 变换可写成

$$
r(z) = p(z) q(z) = \sum_{i,j=0}^{\infty} a_i b_j z^{i+j},
$$

对于 $z \in \mathbf{C}$，$|z| \leqslant \rho$，上式有定义. ■

3.4.2 逆算子

设时间序列$\{Y_n\}$可以由序列$\{X_n\}$通过 L_1 滤波 $\theta(L) = \sum_i \theta_i L^i$ 得到，即 $\forall n$ 有 $Y_n = \theta(L) X_n = \sum_i \theta_i X_{n-i}$. 现考察方程$\{Y_n\} = \theta(L)(\{X_n\})$的可逆性，即是否存在另一个($L_1$?)滤波 $\varphi(L) = \sum_i \varphi_i L^i$ 使得，$\forall n$

$$
X_n = \varphi(L) Y_n = \sum_i \varphi_i Y_{n-i}.
$$

若答案是肯定的，则分别称$\{Y_n\}$和 $\theta(L)$是**可逆的**，称 $\varphi(L) = \theta^{-1}(L)$为逆滤波. 在这种情况下，我们可以由输出$\{Y_n\}$来重构输入$\{X_n\}$. 前面已经介绍了滤波的复合，因此有

$$
\begin{aligned}
\{Y_n\} &= \theta(L)(\{X_n\}) \\
&= \theta(L)(\varphi(L)(\{Y_n\})) \\
&= \theta(L) \cdot \varphi(L)(\{Y_n\}),
\end{aligned}
$$

最后一个等式由 $\theta(L) \circ \varphi(L) = \theta(L) \cdot \varphi(L)$得到. 该方程对任意的序列均成立的充分必要条件为 $\theta(L)\varphi(L) = \mathrm{id} = 1$. 现想导出它们对应的 z 变换的一个必要条件. 假设 $\theta(L)$可逆，其逆算子为 L_1 滤波$\{\varphi(L)\}$，因为 $\theta(L)$和 $\varphi(L)$均为 L_1 滤波，它们对应的 z 变换至少在 $z \in \mathbf{C}$，$|z| \leqslant 1$时是存在的. 因此，它们的积 $\theta(z)\varphi(z)$在单位圆上也存在，从而 $\theta(z)\varphi(z)$在单位圆上恒等于 1，即有 [116]

$$
\theta(z)\varphi(z) = 1, \quad \text{对所有的 } z \in \mathbf{C}, |z| \leqslant 1. \tag{3.6}
$$

因此有

$$
\varphi(z) = \frac{1}{\theta(z)}, \quad \text{对所有的 } |z| \leqslant 1.
$$

所以要得到逆滤波 $\theta^{-1}(L) = \varphi(L) = \sum_i \varphi_i L^i$，需要计算函数 $1/\theta(z)$在单位圆上的级数展开式，然后使用变量变换 $z \mapsto L$. 该方法是可行的，因为(3.6)式隐含了下列引理.

引理 3.4.7 对一个可逆的滤波，其 z 变换 $\theta(z)$在单位圆$\{z \in \mathbf{C}: |z| \leqslant 1\}$上没有根.

▶**例 3.4.8** 考虑下列两个简单例子，一个是线性滤波情形，一个是多项式滤波情形.

（ⅰ）线性情形：考虑 $p(L)=1-aL$. 则当 $|a|<1$ 时，其 z 变换 $p(z)=1-az$ 的根 $1/a$ 在单位圆之外，$p(L)$的逆算子的系数是绝对可求和的. 其逆算子为

$$p^{-1}(L)=\frac{1}{1-aL}=\sum_{i=0}^{\infty}a^iL^i.$$

事实上，其 z 变换为 $p(z)=1-az$，$z\in\mathbf{C}$. 若 $z\neq 1/a$，则

$$p^{-1}(z)=\frac{1}{1-az},$$

且有 $p(z)p^{-1}(z)=1$. 当 $|az|<1$(即 $|z|<1/|a|$)时，$p^{-1}(z)$可以展成下列几何级数：

117

$$p^{-1}(z)=\sum_{i=0}^{\infty}a^iz^i.$$

当 $|a|<1$ 时，级数 $p^{-1}(1)=\sum_i a^i$ 收敛，因此，其收敛半径不小于 1.

（ⅱ）滞后多项式：滤波

$$p(L)=1-a_1L-\cdots-a_pL^p.$$

它有绝对可求和系数的逆算子 $p^{-1}(L)$的充分必要条件是特征多项式

$$p(z)=1-a_1z-\cdots-a_pz^p$$

的根均在单位圆之外，即

$$p(z)=0\quad\Rightarrow\quad |z|>1.$$

下面证明这个结论：因为任意一 n 次多项式 $f(z)=a_0+a_1z+\cdots+a_nz^n$ 可以表示为 $f(z)=a_n\prod_{i=1}^{n}(z-z_i)$，其中 z_1，…，z_n 为多项式的复根，所以特征多项式可以分解为

$$p(z)=-a_p(z-z_1)\cdots(z-z_p),$$

其中 z_1，…，z_p 为 $p(z)$的 p 个复根，且均不等于 0. 将该分解式重写为

$$p(z)=(-a_p)(-1)^p\left(1-\frac{z}{z_1}\right)\cdots\left(1-\frac{z}{z_p}\right)\prod_{j=1}^{p}z_j.$$

记常数 $C=(-a_p)(-1)^p\prod_{j=1}^{p}z_j$，则滞后多项式 $p(L)$是下列线性算子的复合：

$$p_i(L)=1-\frac{L}{z_i},\quad i=1,\cdots,p.$$

将每个因子的 z 变换$\frac{1}{1-z/z_i}$展开成幂级数

$$\frac{1}{1-z/z_i}=\sum_{j=0}^{\infty}\left(\frac{1}{z_i}\right)^jz^j,$$

当 $|z_j|>1$ 时，因每个幂级数的收敛半径大于或等于 1，这些幂级数乘积的收敛半径也大于等于 1，所以得到

118

$$\frac{1}{p(z)}=\frac{1}{-a_p(-1)p_{z_1\cdots z_p}}\prod_{i=1}^{p}\sum_{j=0}^{\infty}\left(\frac{1}{z_j}\right)^jz^j=\sum_{j=0}^{\infty}b_jz^j,$$

其中$\{b_j\}$为系数，这些系数确定了逆滤波 $p^{-1}(L)=\sum_j b_jL^j$. ◀

3.4.3 AR(p)和 AR(∞)过程

本小节介绍自回归过程.

定义 3.4.9

(i) 称随机过程$\{X_t\}$为 **p 阶自回归过程**或 AR(p)过程，如果存在 $p\in\mathbf{N}$，θ_1，…，$\theta_p\in\mathbf{R}$，$\theta_p\neq 0$ 和白噪声过程$\{\varepsilon_t\}$，使得

$$X_t=\theta_1 X_{t-1}+\cdots+\theta_p X_{t-p}+\varepsilon_t,\quad t\in\mathbf{Z}.$$

(ii) 如果存在某个序列$\{\theta_n\}$和一个白噪声过程$\{\varepsilon_t\}$，使得

$$X_t=\sum_{i=1}^{\infty}\theta_i X_{t-i}+\varepsilon_t,\quad t\in\mathbf{Z},$$

那么称$\{X_t\}$为 AR(∞)过程.

(iii) 称$\{X_t:t\in\mathbf{N}_0\}$为 AR(p)过程(X_{-1}，…，X_{-p}已知)，如果

$$X_t=\theta_1 X_{t-1}+\cdots+\theta_p X_{t-p}+\varepsilon_t,\quad t=0,1,\cdots$$

这里$\{\varepsilon_t\}$是白噪声过程，θ_1，…，$\theta_p\in\mathbf{R}$，且 $\theta_p\neq 0$.

注意，在上述定义中，没有要求$\{X_t\}$是平稳的.

▶**例 3.4.10** 考虑 AR(1)方程 $X_t=\rho X_{t-1}+\varepsilon_t$，$t\in\mathbf{Z}$，其中$|\rho|<1$. 由滞后算子的定义，上述方程可以写成 $p(L)X_t=\varepsilon_t$，其中 $p(L)=1-\rho L$，因此，特征多项式可以写成 $p(z)=1-\rho z$. 由上节的结论知，$p(z)$的所有根都在单位圆之外当且仅当$|\rho|<1$ 时有 $p(z)\neq 0$，$\forall z\in\mathbf{C}$，$|z|\leqslant 1$. 因此，可以通过求 $1/p(z)$的级数展开 $\frac{1}{p(z)}=\sum_{i=0}^{\infty}\rho^i z^i$ 来得到 $p(L)$的逆滤波. 于是，线性过程$\{X_t\}$可以写成白噪声的线性组合 [119]

$$X_t=(1-\rho L)^{-1}\varepsilon_t=\sum_{i=0}^{\infty}\rho^i\varepsilon_{t-i}.\tag{3.7}$$

由以上写法可以得到 X_t 的自协方差表达式为

$$\begin{aligned}\gamma_X(h)&=\mathrm{Cov}\Big(\sum_{i=0}^{\infty}\rho^i\varepsilon_{t+h-i},\sum_{j=0}^{\infty}\rho^j\varepsilon_{t-j}\Big)\\&=\sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\rho^{i+j}\gamma_\varepsilon(h-i+j).\end{aligned}$$

注意到 $\gamma_\varepsilon(h-i+j)=\sigma^2$ 当且仅当 $i=h+j$，否则为 0. 所以

$$\begin{aligned}\gamma_X(h)&=\sigma^2\sum_{j=0}^{\infty}\rho^{2j}\rho^h\\&=\frac{\rho^h}{1-\rho^2}\sigma^2.\end{aligned}$$

由上式可以看出，自协方差以指数速率 $\alpha(h)\sim \mathrm{e}^{-h\gamma}$ 衰减，其中 $\gamma=-\log(\rho)>0$. ◀

▶**例 3.4.11** 考虑零初值的 AR(1)过程，即

$$X_0=0,\quad X_t=\rho X_{t-1}+\varepsilon_t,\quad t\geqslant 1,\quad \varepsilon_t\sim \mathrm{WN}(0,\sigma^2),$$

上述过程并不是平稳的，但当 $t\to\infty$ 时收敛于一个平稳过程. 事实上，因为

$$X_0 = 0, X_1 = \varepsilon_1, X_2 = \rho\varepsilon_1 + \varepsilon_2,$$

一般地，有

$$X_t = \sum_{i=1}^{t} \rho^{t-i}\varepsilon_i = \sum_{i=0}^{t-1} \rho^i \varepsilon_{t-i}, \tag{3.8}$$

随着 t 的增大，上述方程不断地接近 AR(1)方程的平稳解 $\sum_{i=0}^{\infty}\rho^i\varepsilon_{t-i}$，其中 L_2 距离为

$$\sqrt{E\Big(X_t - \sum_{i=0}^{\infty}\rho^i\varepsilon_{t-i}\Big)^2} = \sigma\sqrt{\sum_{i\geqslant t}\rho^{2i}}.$$

方程(3.8)的方差为

120

$$\sigma_t^2 = \mathrm{Var}(X_t) = \sigma^2\sum_{i=0}^{t-1}\rho^{2i},$$

可以看出，方程依赖于 t. 显然，我们有

$$\sigma_t^2 \to \frac{\sigma^2}{1-\rho^2}, \quad t\to\infty.$$

而且对 $h\geqslant 0$，有

$$\begin{aligned}\mathrm{Cov}(X_t, X_{t+h}) &= \mathrm{Cov}\Big(\sum_{i=1}^{t}\rho^{t-i}\varepsilon_i, \sum_{j=1}^{t+h}\rho^{t+h-j}\varepsilon_j\Big)\\ &= \sigma^2 p^h\sum_{i=1}^{t}\rho^{2(t-i)}.\end{aligned}$$

所以它的自协方差也依赖于 t，但它是收敛的，即当 $t\to\infty$时，

$$\mathrm{Cov}(X_t, X_{t+h}) \to \frac{\sigma^2\rho^h}{1-\rho^2}.$$

当 t 充分大时，(3.8)式和平稳解(3.7)的差别是非常小的，往往可以忽略不计. ◀

下面给出一个 AR(p)过程可以表示为 MA(∞)过程的条件.

命题 3.4.12 设

$$X_t = \sum_{i=1}^{p}\theta_i X_{t-i} + \varepsilon_t, \quad t\in\mathbf{Z}, \quad \varepsilon_t \sim \mathrm{WN}(0,\sigma^2)$$

是一个 AR(p)过程，如果滞后多项式 $p(z)=1-\theta_1 z-\cdots-\theta_p z^p\ (z\in\mathbf{C})$的所有的根均在单位圆之外，那么$\{X_t\}$可以表示为平稳的 MA($\infty$)过程 $X_t = \sum_{i=0}^{\infty}c_i\varepsilon_{t-i}$.

证明 注意到 AR(p)过程等价于

$$\varepsilon_t = X_t - \sum_{i=1}^{p}\theta_i X_{t-i} = p(L)X_t, \quad t\in\mathbf{Z}.$$

121

如果 $p(z)=0\Rightarrow|z|>1$，方程是可逆的，那么有 $X_t = p(L)^{-1}\varepsilon_t = \sum_{i=0}^{\infty}c_i\varepsilon_{t-i}$，其中系数 c_i 由方程 $\sum_{i=0}^{\infty}c_i z^i \overset{!}{=} \frac{1}{p(z)}$ 确定，且有 $\sum_i|c_i|<\infty$，这意味着 X_t 可由 L_1 滤波作用于 ε_t 得到，因此是平稳的. ■

一般地，满足表达式 $X_t = q(L)\varepsilon_t = \sum_{i=0}^{\infty} c_i \varepsilon_{t-i}$ 的 MA(∞)过程的系数 c_i 可以通过下列方法计算：首先由 $p(L)=1-\theta_1 L-\cdots-\theta_p L^p$，有 $q(L)p(L)=1$，它等价于

$$(c_0 + c_1 L + c_2 L^2 + \cdots)(1 - \theta_1 L - \cdots - \theta_p L^P) = 1.$$

显然，等式左边为

$$c_0 + (c_1 - \theta_1 c_0)L + (c_2 - \theta_1 c_1 - \theta_2 c_0)L^2 + \cdots + (c_i - \theta_1 c_{i-1} - \cdots - \theta_p c_{i-p})L^i + \cdots$$

通过比较系数，可以建立下列方程组

$$c_0 = 1$$
$$c_1 = \theta_1 c_0 = \theta_1$$
$$c_2 = \theta_1 c_1 + \theta_2 c_0 = \theta_1^2 + \theta_2.$$

或者使用另外一种记法，设 $c_{-p}=\cdots=c_{-1}=0$，对 $i\geqslant 1$ 必须解下列差分方程

$$c_0 = 1,\quad c_i - \theta_1 c_{i-1} - \cdots - \theta_p c_{i-p} = 0.$$

3.4.4 ARMA 过程

定义 3.4.13(ARMA(p, q)过程)

(ⅰ) 称$\{X_t: t\in\mathbf{Z}\}$为 ARMA(p, q)过程，p, $q\in\mathbf{N}$，如果它是平稳的且满足

$$X_t - \phi_1 X_{t-1} - \cdots - \phi_p X_{t-p} = \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_q \varepsilon_{t-q},\quad t \in \mathbf{Z},$$

其中 ϕ_1, …, ϕ_p, θ_1, …, $\theta_q\in\mathbf{R}$ 是常数，且 $\phi_p\neq 0$, $\theta_q\neq 0$, $\{\varepsilon_t\}$是白噪声过程.

(ⅱ) 如果 $X_t-\mu\sim$ ARMA(p, q)，那么称$\{X_t: t\in\mathbf{Z}\}$是均值为 μ 的 ARMA(p, q)过程.

(ⅲ) 称 ARMA(p, q)过程是因果的(关于 $Z_t\sim$ WN(0, σ^2))，如果存在滤波$\{\psi_k: k\in\mathbf{N}_0\}$使得

$$X_t = \sum_{i=0}^{\infty} \psi_i Z_{t-i}.$$

122

对 $z\in\mathbf{C}$，定义下列滞后多项式

$$\phi(z) = 1 - \phi_1 z - \cdots - \phi_p z^p,$$
$$\theta(z) = 1 + \theta_1 z + \cdots + \theta_q z^q,$$

称 ϕ 为 **AR 滞后多项式**，θ 为 **MA 滞后多项式**. 于是 ARMA(p, q)方程可写为 $\phi(L)X_t=\theta(L)\varepsilon_t$的形式. 在以下内容中，均假设滞后多项式没有公共根.

定理 3.4.14 设$\{X_t\}$是一个平稳 ARMA(p, q)过程，滞后多项式分别为$\phi(z)$和$\theta(z)$，即 $\phi(L)X_t=\theta(L)\varepsilon_t$. 设 $\phi(z)$的根均在单位圆外，那么 $X_t=\psi(L)\varepsilon_t$ 是平稳的 ARMA(p, q)因果过程，其系数由 $\psi(z) = \sum_{i=0}^{\infty}\psi_i z^i$ 给出，对$|z|\leqslant 1$,

$$\psi(z) = \frac{1 + \theta_1 z + \cdots + \theta_q z^q}{1 - \phi_1 z - \cdots - \phi_p z^p}.$$

证明 对一个白噪声 ε_t，序列 $Y_t=\theta(L)\varepsilon_t$ 是平稳的. 只要 $\phi(z)$的根满足$|z|>1$，则可以解出 $\phi(L)X_t=Y_t$. ■

下面的经典例子有助于理解上述结论.

▶**例 3.4.15** 考虑下列 ARMA(1，1)过程，

$$X_t - \phi X_{t-1} = \varepsilon_t + \theta\varepsilon_{t-1},$$

其中系数 $\phi \neq -\theta$ 保证了滞后多项式没有公共根. 为了确定系数 $\psi_i(i \geqslant 0)$，考虑方程

123

$$\sum_{i=0}^{\infty} \psi_i z^i = \psi(z) = \frac{\theta(z)}{\phi(z)} = \frac{1+\theta z}{1-\phi z}.$$

注意到

$$\begin{aligned}\frac{1+\theta z}{1-\phi z} &= (1+\theta z)\sum_{i=0}^{\infty}(\phi z)^i \\ &= (1+\theta z)\Big(1+\sum_{i=0}^{\infty}(\phi z)^i\Big) \\ &= 1+\sum_{i=0}^{\infty}(\phi+\theta)\phi^{i-1}z^i,\end{aligned}$$

因此，$\psi_0 = 1$ 且对 $i \geqslant 1$，有 $\psi_i = (\phi+\theta)\phi^{i-1}$. ◀

3.5 频域分析

3.5.1 频谱

设$\{X_t\}$为一个平稳过程，且

$$\mu = E(X_t),$$
$$\gamma_h = \mathrm{Cov}(X_1, X_{1+h}), \quad h \in \mathbf{Z}.$$

若函数 $g_X(z) = \sum_{k=-\infty}^{\infty} \gamma_k z^k$ 存在，则称它为**自协方差母函数**. g_X 在单位圆上有定义的充分条件是自协方差满足绝对可求和性. 注意，对 $k \in \mathbf{N}_0$ 有 $\gamma_k = g_X^{(k)}(0)/k!$

在以下的内容中，将用到虚数 i，即 $\mathrm{i}^2 = -1$.

定义 3.5.1 设$\{X_t\}$为一个平稳过程，自协方差函数$\{\gamma_k\}$满足 $\sum_k |\gamma_k| < \infty$，则称

$$f_X(\omega) = \frac{1}{2\pi} g_X(\mathrm{e}^{-\mathrm{i}\omega}) = \frac{1}{2\pi}\sum_{k=-\infty}^{\infty} \gamma_k \mathrm{e}^{-\mathrm{i}\omega k}, \quad \omega \in \mathbf{R},$$

为$\{X_t\}$的**频谱**或**谱密度**.

在时间序列分析中，谱密度被广泛地记为 f_X. 为了不与概率密度混淆，本书将其记为 $f_X(\omega)$. 根据欧拉公式 $\mathrm{e}^{\mathrm{i}z} = \cos(z) + \mathrm{i}\sin(z)$，可以直接得到下列结果：

124

$$\mathrm{e}^{-\mathrm{i}\omega k} = \cos(\omega k) - \mathrm{i}\sin(\omega k), \quad \mathrm{e}^{-\mathrm{i}\omega k} + \mathrm{e}^{\mathrm{i}\omega k} = 2\cos(\omega k),$$

再由自协方差的对称性，即 $\gamma_k = \gamma_{-k}$，则有下列经常使用的公式：

$$f_X(\omega) = \frac{1}{2\pi}\Big(\gamma_0 + 2\sum_{k=1}^{\infty}\gamma_k\cos(\omega k)\Big) = \frac{1}{2\pi}\sum_{k=-\infty}^{\infty}\gamma_k\cos(\omega k).$$

由于 $f_X(\omega)$是周期为 2π 的周期函数且是偶函数，即 $f_X(-\omega)=f_X(\omega)$，所以只需在区间 $\omega\in[-\pi,\pi]$上研究其性质即可.

值得一提的是，因为 $\int_{-\pi}^{\pi}\cos(\omega k)\mathrm{d}\omega = 2\pi\mathbf{1}(k=0)$，所以有

$$\int_{-\pi}^{\pi} f_X(\omega)\mathrm{d}\omega = \frac{1}{2\pi}\sum_k \gamma_k \int_{-\pi}^{\pi}\cos(\omega k)\mathrm{d}\omega = \gamma_0 = \mathrm{Var}(X_t),$$

上式即表明，对谱密度在$[-\pi,\pi]$上积分，就得到过程的边际方差.

对于线性过程，下面给出一个用 z 变换计算谱密度的漂亮公式.

引理 3.5.2 对于线性过程 $X_t = \sum_{k=0}^{\infty}\psi_k\varepsilon_{t-k}, \varepsilon_t \sim \mathrm{WN}(0,\sigma^2)$,

$$g_X(z) = \sigma^2\psi(z)\psi(z^{-1}),$$

其谱密度为

$$f_X(\omega) = \frac{\sigma^2}{2\pi}\psi(\mathrm{e}^{-\mathrm{i}\omega})\psi(\mathrm{e}^{\mathrm{i}\omega}),$$

其中 $\psi(z)$表示 z 变换. 更一般地，有 $g_X(z)=\psi(z^{-1})g_\varepsilon(z)\psi(z)$成立.

证明 为了计算 $g_X(z) = \sum_k \gamma_k z^k$，插入公式

$$\gamma_X(k) = \sum_i\sum_j \psi_i\psi_j\gamma_\varepsilon(k+i-j),$$

则有

$$g_X(z) = \sum_k\sum_i\sum_j \psi_i\psi_j\gamma_\varepsilon(k+i-j)z^k.$$

使用变量替换 $h=k+i-j$，使得 $z^k=z^h z^{-i}z^j$，则得到

$$g_X(z) = \Big(\sum_i \psi_i z^{-i}\Big)\Big(\sum_h \gamma_\varepsilon(h)z^h\Big)\Big(\sum_j \psi_j z^j\Big),$$

上式等于 $\psi(z^{-1})g_\varepsilon(z)\psi(z)$，其中，对白噪声过程$\{\varepsilon_t\}$，$g_\varepsilon(z)=\sigma^2$，因此结论成立. ■

注 3.5.3 注意，$f_X(\omega)=\frac{\sigma^2}{2\pi}|\psi(\mathrm{e}^{-\mathrm{i}\omega})|^2$，这是因为有$\overline{\psi(z)}=\psi(\bar{z})$和$\overline{\mathrm{e}^{\mathrm{i}w}}=\mathrm{e}^{-\mathrm{i}w}$成立. 125

下面的例子同样是经典的，它也有助于理解上述结论.

例 3.5.4 考虑下列平稳 AR(1)过程

$$X_t = \phi X_{t-1} + \varepsilon_t, \quad \varepsilon_t \sim \mathrm{WN}(0,\sigma^2).$$

则其谱密度 $f_X(\omega)$为

$$f_X(\omega) = \frac{\sigma^2}{2\pi}\frac{1}{1-2\phi\cos(\omega)+\phi^2}.$$

事实上，我们有

$$\begin{aligned} f_X(\omega) &= \frac{\sigma^2}{2\pi}\frac{1}{1-\phi\mathrm{e}^{-\mathrm{i}\omega}}\frac{1}{1-\phi\mathrm{e}^{\mathrm{i}\omega}} \\ &= \frac{\sigma^2}{2\pi}\frac{1}{|1-\phi\mathrm{e}^{\mathrm{i}\omega}|^2} \\ &= \frac{\sigma^2}{2\pi}\frac{1}{|1-\phi\cos(\omega)-\mathrm{i}\sin(\omega)|^2} \end{aligned}$$

$$
= \frac{1}{(1-\phi\cos(\omega))^2+\phi^2\sin^2(\omega)}
$$

$$
= \frac{\sigma^2}{2\pi}\frac{1}{1-2\phi\cos(\omega)+\phi^2}.
$$

考虑下列 ARMA(p, q)过程

$$
\phi(L)X_t=\theta(L)\varepsilon_t,\quad \varepsilon_t\sim \mathrm{WN}(0,\sigma^2),
$$

其中滞后多项式为 $\phi(z)=1-\phi_1 z-\cdots-\phi_p z^p$ 和 $\theta(z)=1+\theta_1 z+\cdots+\theta_q z^q$，因为，$X_t=\left.\frac{\theta(z)}{\phi(z)}\right|_{z=L}\varepsilon_t$，则频谱为

$$
f_X(\omega)=\left.\frac{\sigma^2}{2\pi}\frac{\theta(z)}{\phi(z)}\frac{\theta(z^{-1})}{\phi(z^{-1})}\right|_{z=\mathrm{e}^{-\mathrm{i}\omega}}
$$

代入推得频谱为

$$
\frac{\sigma^2}{2\pi}\frac{1+\theta_1\mathrm{e}^{-\mathrm{i}\omega}+\cdots+\theta_q\mathrm{e}^{-\mathrm{i}q\omega}}{1-\phi_1\mathrm{e}^{-\mathrm{i}\omega}-\cdots-\phi_p\mathrm{e}^{-\mathrm{i}p\omega}}\frac{1+\theta_1\mathrm{e}^{\mathrm{i}\omega}+\cdots+\theta_q\mathrm{e}^{\mathrm{i}q\omega}}{1-\phi_1\mathrm{e}^{\mathrm{i}\omega}-\cdots-\phi_p\mathrm{e}^{\mathrm{i}p\omega}}.
$$

3.5.2 周期图法

为了得到频谱估计量，自然会想到用某个样本自协方差代替自协方差，比如

126

$$
\hat{\gamma}_T(h)=\frac{1}{T}\sum_{t=1}^{T-h}(X_t-\overline{X}_T)(X_{t+h}-\overline{X}_T),\quad \hat{\gamma}_T(-h)=\hat{\gamma}_T(h),
$$

或

$$
\widetilde{\gamma}_T(h)=\frac{1}{T}\sum_{t=1}^{T-h}X_tX_{t+h},\quad \widetilde{\gamma}_T(-h)=\widetilde{\gamma}_T(h),
$$

其中 $\overline{X}_T=\frac{1}{T}\sum_{t=1}^{T}X_t$.

定义 3.5.5 称随机函数

$$
I_T(\omega)=\frac{1}{2\pi}\sum_{h=-(T-1)}^{T-1}\widetilde{\gamma}_T(h)\mathrm{e}^{-\mathrm{i}\omega h}
$$

为**周期图**，称

$$
J_T(\omega)=\frac{1}{2\pi}\sum_{h=-(T-1)}^{T-1}\hat{\gamma}_T(h)\mathrm{e}^{-\mathrm{i}\omega h}
$$

为**中心周期图**.

引理 3.5.6 周期图和中心周期图满足下列等式：

$$
I_T(\omega)=\frac{1}{2\pi T}\left|\sum_{t=1}^{T}X_t\mathrm{e}^{-\mathrm{i}\omega t}\right|^2,
$$

$$
J_T(\omega)=\frac{1}{2\pi T}\left|\sum_{t=1}^{T}(X_t-\overline{X}_T)\mathrm{e}^{-\mathrm{i}\omega t}\right|^2.
$$

下面引入赫格罗兹(Herglotz)定理，它是一个非常重要的定理. 该定理表明，自协方差函数可以由频谱确定.

定理 3.5.7 设$\{X_t\}$是一个平稳过程，其自协方差函数$\{\gamma_k\}$满足$\sum_k|\gamma_k|<\infty$，则

（ⅰ）$I_T(\omega)$是谱密度$f_X(\omega)$的渐近无偏估计，即有

$$\lim_{T\to\infty}E(I_T(\omega))=f_X(\omega).$$

（ⅱ）谱密度$f_X(\omega)$是非负连续的偶函数，并且可以通过下面的变换来得到自协方差函数：

$$\gamma_h=\int_{-\pi}^{\pi}\cos(\omega h)f_X(\omega)\mathrm{d}\omega,\quad h\in\mathbf{Z}.$$

证明 要证谱密度

$$f(\omega)=f_X(\omega)=\frac{1}{2\pi}\sum_t\gamma_t\cos(\omega t)\tag{3.9}$$

127

的连续性，只需证明右边级数是一致收敛即可．对所有的t和ω，由$|\gamma_t\cos(\omega t)|\leqslant|\gamma_t|$知，$|f_X(\omega)|\leqslant\sum_t|\gamma_k|<\infty$. 回忆$L_2$空间内积的定义，对$f$，$g\in L_2([-\pi,\pi];\lambda)$，定义内积$(f,g)=\int_{-\pi}^{\pi}f(x)g(x)\mathrm{d}x$，则函数

$$\varphi_0(t)=\frac{1}{\sqrt{2\pi}},$$

$$\varphi_k(t)=\frac{1}{\sqrt{\pi}}\cos(kt),\quad k\in\mathbf{N},$$

关于上述内积构成一个正交系．因为f在$[-\pi,\pi]$上连续且$f(-\pi)=f(\pi)$，则傅立叶级数$f(x)=\sum_{k=0}^{\infty}(f,\varphi_k)\varphi_k(x)$是一致收敛的．这意味着有

$$\begin{aligned}f(x)&=\frac{1}{2\pi}\int_{-\pi}^{\pi}f(t)\mathrm{d}t+2\sum_{k=1}^{\infty}\frac{1}{2\pi}\int_{-\pi}^{\pi}f(t)\cos(kt)\mathrm{d}t\cdot\cos(kx)\\&=\sum_{k\in\mathbf{Z}}\frac{1}{2\pi}\int_{-\pi}^{\pi}f(t)\mathrm{d}t\cos(kx).\end{aligned}$$

将该等式与(3.9)式比较，由傅立叶级数展开式中系数的唯一性知，

$$\gamma_k=\int_{-\pi}^{\pi}f(t)\cos(kt)\mathrm{d}t$$

对所有的k成立．

进一步地，由引理 3.5.6 知$I_T(\omega)$是非负的，从而$E(I_T(\omega))\geqslant0$. 现详细地来讨论其期望，由线性性有

$$\begin{aligned}E(I_T(\omega))&=\frac{1}{2\pi T}\sum_{s,t=1}^{T}E(X_sX_t)\mathrm{e}^{\mathrm{i}\omega(s-t)}\\&=\frac{1}{2\pi T}\sum_{s,t=1}^{T}\gamma_{|s-t|}\mathrm{e}^{\mathrm{i}\omega(s-t)}.\end{aligned}$$

它是对称矩阵的T^2个元素的和，元素为$\gamma_{|s-t|}\mathrm{e}^{\mathrm{i}\omega(s-t)}$. 因为

$$\mathrm{e}^{\mathrm{i}\omega(s-t)}+\mathrm{e}^{\mathrm{i}\omega(t-s)}=2\cos(\omega(s-t)),$$

记矩阵元素为$a_{st}=\gamma_{|s-t|}\mathrm{e}^{\mathrm{i}\omega(s-t)}$，则有$a_{st}+a_{ts}=2\gamma_{|s-t|}\cos(\omega(|s-t|))$，利用上述性质，

128

对矩阵所有的对角线元素求和，则得到

$$E(I_T(\omega)) = \frac{1}{2\pi}\sum_{|t|<T}\left(1-\frac{|t|}{T}\right)\gamma_t\cos(\omega t).$$

上述结果可以写成关于 $\mathbf{Z}$ 上的计数测度 $\mathrm{d}\nu(t)$ 的积分[⊖]，即 $\int F_T(t;\omega)\mathrm{d}\nu(t)$，其中，

$$F_T(t;\omega) = \begin{cases} \frac{1}{2\pi}\left(1-\frac{|t|}{T}\right)\gamma_t\cos(\omega t), & |t| \leqslant T-1, \\ 0, & |t| > T-1. \end{cases}$$

则得到了

$$0 \leqslant E(I_T(\omega)) = \int F_T(t;\omega)\mathrm{d}\nu(t).$$

下面使用控制收敛定理证明 $E(I_T(\omega))$ 的极限也是非负的. 首先注意到 $|F_T(t;\omega)| \leqslant |\gamma_t|$ 对所有的 $t\in\mathbf{Z}$ 成立，且由假设知 $\int|\gamma_t|\mathrm{d}\nu(t) = \sum_{t\in\mathbf{Z}}|\gamma_t| < \infty$. 因此，$t\mapsto|\gamma_t|\mathbf{1}_{\mathbf{Z}}(t)$ 是一个 $\mathrm{d}\nu$ 可积的控制函数. 且 $F_T(t;\omega)$ 处处收敛，即当 $T\to\infty$ 时，有

$$F_T(t;\omega) = \frac{1}{2\pi}\left(1-\frac{|t|}{T}\right)\cos(\omega t) \to \frac{1}{2\pi}\gamma_t\cos(\omega t),$$

从而当 $T\to\infty$ 时有

$$\begin{aligned} 0 &\leqslant \int F_T(t;\omega)\mathrm{d}\nu(t) \\ &\to \int \frac{1}{2\pi}\gamma_t\cos(\omega t)\mathrm{d}\nu(t) \\ &= \frac{1}{2\pi}\sum_{t=-\infty}^{\infty}\gamma_t\cos(\omega t) \\ &= f_X(\omega). \end{aligned}$$

■

我们已经证明了谱密度 $f_X(\omega)$ 是非负的，且满足 $\int_{-\pi}^{\pi}f_X(\omega)\mathrm{d}\omega = \mathrm{Var}(X_t) < \infty$. 因此，它定义了一个有限测度.

129

定义 3.5.8 称定义在 $((-\pi,\pi],\mathcal{B}(-\pi,\pi])$ 上的 Borel 测度

$$\nu_X(A) = \int_A f_X(\omega)\mathrm{d}\omega,\quad A\subset(-\pi,\pi]$$

为 $\{X_t\}$（或 $f_X(\omega)$）的**谱测度**. 对应的（广义）分布函数

$$F_X(\omega) = \int_{-\pi}^{\omega} f_X(\lambda)\mathrm{d}\lambda,\quad \omega\in(-\pi,\pi],$$

称为**谱分布函数**.

只要给定了自协方差函数（过程不必是平稳的），就可以定义一个谱密度，同样，给定

⊖ 对计数测度 $\mathrm{d}\nu$ 和实值函数 f，只要级数 $\sum_{n\in\mathbf{Z}}f(n)$ 存在，就有 $\int f(x)\mathrm{d}\nu(x) = \sum_{n\in\mathbf{Z}}f(n)$.

了自协方差函数也可以定义谱测度和谱分布函数.

注 3.5.9 对取值于 $\mathbf{C}$ 的过程 $\{X_t\}$，有下列结论：

$$\gamma_X(h) = \int_{-\pi}^{\pi} \mathrm{e}^{\mathrm{i}h\omega} f_X(\omega)\mathrm{d}\omega,$$

等价地，有

$$\gamma_X(h) = \int_{-\pi}^{\pi} \mathrm{e}^{\mathrm{i}h\omega} \mathrm{d}F_X(\omega) = \int_{-\pi}^{\pi} \mathrm{e}^{\mathrm{i}h\omega} \mathrm{d}\nu_X(\omega).$$

可以将上述结论推广到只要求谱密度存在，不要求自协方差可求和的情形，从而简化论证，证明也更简洁. 下面给出一个一般性的结论.

定理 3.5.10（赫格罗兹引理）

（ⅰ）设 $\gamma:\mathbf{Z}\to\mathbf{C}$ 是一个半正定的函数，则存在唯一的谱测度 ν，使得

$$\gamma(h) = \int_{(-\pi,\pi]} \mathrm{e}^{\mathrm{i}h\omega} \mathrm{d}\nu(\omega), \quad h \in \mathbf{Z}. \tag{3.10}$$

（ⅱ）若 ν 是定义在 $(-\pi, \pi]$ 的 Borel σ 代数上的任一有限测度，那么(3.10)式给出的函数是 $\mathbf{Z}$ 上的一个复值半正定函数，从而它是一个平稳过程的自协方差函数.

注 3.5.11 设过程 $\{X_t\}$ 的谱分布函数为 $F_X(\omega)$，自协方差为 $\gamma_X(k)$，则有

$$F_X(\pi) = \int_{-\pi}^{\pi} f(\omega)\mathrm{d}\omega = \gamma_X(0) = \mathrm{Var}(X_t).$$

130

因为 $\mathrm{Var}(X_t)<\infty$，

$$G_X(\omega) = F_X(\omega)/F_X(\pi), \quad \omega \in [-\pi,\pi],$$

则当 $\omega<-\pi$ 时，$G_X(\omega)=0$，当 $\omega>\pi$ 时，$G_X(\omega)=1$，因此上式定义了概率测度在 $[-\pi, \pi]$ 上的分布函数. 它和自相关系数的关系为：

$$\rho_X(h) = \frac{\gamma_X(h)}{\gamma_X(0)} = \int_{-\pi}^{\pi} \mathrm{e}^{\mathrm{i}h\omega} \mathrm{d}G_X(\omega), \quad h \in \mathbf{Z}.$$

如果过程 $\{X_t\}$ 的谱密度不存在，那么可以研究其谱分布函数. 特别地，若 $\gamma_X(h) = \sum_j \alpha_j \mathrm{e}^{\mathrm{i}\omega_j h}$，其中系数 $\alpha_j \in \mathbf{R}$，$\omega_j \in [-\pi, \pi]$ 是固定频率. 则由 $\gamma_X(h) = \int_{-\pi}^{\pi} \mathrm{e}^{\mathrm{i}\omega h} \mathrm{d}F_X(\omega)$，得到分布函数

$$F_X(\omega) = \sum_j \alpha_j \delta_{\omega_j}(\omega),$$

其中，δ_a 表示点 a 处的单点测度（即狄拉克测度），谱分布函数只在可数集 $\{\omega_j\}$ 上取值，且在频率 ω_j 的权重为 α_j.

下面的例子说明了两点：第一，上述的例子并不是偶然的；第二，称 ω_j 为频率是有其物理意义的.

▶**例 3.5.12** 考虑下列的随机周期函数

$$X_t = A\cos(\omega t) + B\sin(\omega t), \quad t \in \mathbf{Z},$$

其中 $\omega\in(-\pi, \pi]$ 为一个固定频率，A 和 B 为随机振幅且互相独立. 假设有 $E(A)=E(B)=0$，且 $\mathrm{Var}(A)=\mathrm{Var}(B)=\sigma^2\in(0, \infty)$. 由直接计算可得

$$\gamma_X(h) = \sigma^2 \cos(\omega h) = \sigma^2 \frac{e^{i\omega h} + e^{-i\omega h}}{2}.$$

因此，谱分布函数为

$$F_X(\lambda) = \begin{cases} 0 & \lambda < -\omega, \\ \sigma^2/2, & \lambda \in [-\omega, \omega), \\ \sigma^2, & \lambda \geqslant \omega. \end{cases}$$

131 F_X 在点 $-\omega$ 和 ω 上都取值 $\sigma^2/2$. ◀

3.6 ARMA 过程的估计

ARMA 时间序列的估计和推断一般是建立在极大似然估计基础之上的，它需要假设驱动该过程的白噪声序列是服从正态分布的独立同分布随机变量序列.

设有一个均值为 μ 的 ARMA(p, q)平稳过程，满足

$$X_t - \mu - \phi_1(X_{t-1} - \mu) - \cdots - \phi_p(X_{t-p} - \mu) = \theta_1 \varepsilon_{t-1} + \cdots + \theta_q \varepsilon_{t-q} + \varepsilon_t,$$

等价于

$$\phi(L)(X_t - \mu) = \theta(L)\varepsilon_t,$$

其中，滞后多项式 $\phi(L) = \sum_{i=0}^{p} \phi_i L^i$，$\theta(L) = \sum_{i=0}^{q} \theta_i L^i$，且有 $\phi_0 = \theta_0 = 1$. 假定 $\phi(z)$在单位圆内无根，则过程是可逆的，可以得到线性过程 X_t 的表达式 $X_t = \mu + \psi(L)\varepsilon_t$，其中 $\psi(L) = \sum_{i=0}^{\infty} \psi_i L^i$ 满足 $\psi(z) = \theta(z)/\phi(z) = \sum_{i=0}^{\infty} \psi_i z^i$，$|z| \leqslant 1$.

因为该线性过程也是高斯过程，则有限序列 X_1, …, X_T 是多元正态的，即满足

$$(X_1, \cdots, X_T) \sim N(\mu, \sigma^2 \Sigma(\vartheta)),$$

这里的协方差矩阵 $\Sigma(\vartheta)$依赖于参数向量 $\vartheta = (\phi_1, \cdots, \phi_p, \theta_1, \cdots, \theta_q)$. 那么(精确的)似然函数为

$$L(\vartheta, \mu, \sigma^2 \mid x) = \left(\frac{1}{2\pi\sigma^2}\right)^{T/2} \det(\Sigma)^{-1/2} \exp\left\{-\frac{(x - \mu\mathbf{1})'\Sigma(\vartheta)^{-1}(x - \mu\mathbf{1})}{2\sigma^2}\right\},$$

其中 $x = (x_1, \cdots, x_T)$是 X_1, …, X_T 的观测值，$\mathbf{1} = (1, \cdots, 1)' \in \mathbf{R}^d$. 现已有很多计算 $L(\vartheta, \mu, \sigma^2 \mid X = x)$或对数似然函数的高效算法，通过在那些参数值定义的参数集上最大化似然函数，可以得到 ARMA(p, q)过程的平稳解.

对 AR(p)模型，通常考虑的是它的条件似然函数. 对于样本 X_{-p+1}, …, X_0, X_1, …, X_T，运用条件概率法则 $f_{(X,Y)} = f_{Y \mid X} f_X$ 因式分解其精确似然，有

$$\begin{aligned} f_{(X_{-p+1}, \cdots, X_T)} &= f_{X_T \mid (X_{-p+1}, \cdots, X_{T-1})} f_{(X_{-p+1}, \cdots, X_{T-1})} \\ &= \cdots \\ &= \prod_{t=1}^{T} f_{X_t \mid (X_{t-p+1}, \cdots, X_{t-1})} f_{(X_{-p+1}, \cdots, X_0)}. \end{aligned}$$

132

若给定 X_{-p+1}, …, X_0，最后一个因子可以忽略. 因此**条件似然**函数为

$$L_c(\vartheta|x_1,\cdots,x_T)=\prod_{t=1}^{T}f_{X_t|(X_{t-p},\cdots,X_{t-1})}(x_t|x_{t-p},\cdots,x_{t-1};\vartheta),$$

其中 $\vartheta=(\theta_1, \cdots, \theta_p)$，使上述函数最大的参数值称为极大似然估计量. 特别地，对于均值 $\mu=0$ 的 AR(p)模型，有 $X_t-\sum_{j=1}^{p}\theta_j X_{t-j}\sim N(0,\sigma^2)$，可以推出

$$L_c(\vartheta|X_1,\cdots,X_T)=\left(\frac{1}{2\pi\sigma^2}\right)^{T/2}\exp\left\{-\frac{1}{2\sigma^2}\sum_{t=1}^{T}\left(X_t-\sum_{j=1}^{p}\theta_j X_{t-j}\right)^2\right\}.$$

则使最小二乘方判据

$$Q(\vartheta)=\sum_{t=1}^{T}\left(X_t-\sum_{j=1}^{p}\theta_j X_{t-j}\right)^2$$

达到最小的参数值就为条件极大似然估计量，它是下列线性方程组的解：

$$\widetilde{\Gamma}_T\vartheta+\widetilde{g}_T=0,$$

其中

$$\widetilde{\Gamma}_T=\left(\frac{1}{T}\sum_{t=1}^{T}X_{t-r}X_{t-s}\right)_{r=1,\cdots,p;s=1,\cdots,p},\quad \widetilde{g}_T=\left(\frac{1}{T}\sum_{t=1}^{T}X_tX_{t-r}\right)_{r=1}^{p}.$$

上述方程组表明，条件最小二乘估计量 $\hat{\vartheta}_T$ 依赖于时间序列的样本自协方差.

3.7 (G)ARCH 模型

前面讨论了相依结构和条件一阶矩模型. 下面要介绍一类著名的模型，目标是对条件波动率建模. 这项杰出成果归功于 Robert Engle，因此他获得了 2003 年的诺贝尔经济学奖，该奖是为了纪念 Alfred Nobel 而设立的.

在这类模型里，白噪声过程有鞅差分结构. 在介绍模型之前，首先证明：任意一个 L_2 鞅可以表示成两个因子的乘积，其中一个为基于历史信息的条件波动率. 因此，这类模型可看成是简化的条件波动率参数模型.

设 $\{X_t:t\in\mathbf{Z}\}$ 是一个 L_2 鞅差序列，即 $E|X_t|^2<\infty$ 和 $E(X_t|\mathcal{F}_{t-1})=0$ 对所有的 t 成立，其中 $\mathcal{F}_t=\sigma(X_s:s\leqslant t)$ 是自然滤子.

133

定义 3.7.1 给定历史信息 $\mathcal{F}_{t-1}$，称条件期望

$$\sigma_t^2=E(X_t^2|\mathcal{F}_{t-1}),\quad t\in\mathbf{Z},\tag{3.11}$$

为**条件方差**，而 $\sigma_t=\sqrt{\sigma_t^2}$ 称为**条件波动率**.

注意到 $\sigma_t^2=H_t(X_{t-1}, X_{t-2}, \cdots)$ 是一个 Borel 可测函数 H_t. 在下面的内容中，设对常数 c_t 所有的 t，$h_t\geqslant c_t>0$ 几乎必然成立. 记 $u_t=\dfrac{X_t}{\sigma_t}$，$t\in\mathbf{Z}$，由定义有 $X_t=\sigma_t u_t$，$t\in\mathbf{Z}$. 容易验证下列结论.

引理 3.7.2 在假定(3.11)下，$\{u_t\}$ 是一个 $\mathcal{F}_t$ 鞅差序列，且满足 $\mathrm{Var}(u_t|\mathcal{F}_{t-1})=1$.

从而 $\{u_t\}$ 也是一个白噪声过程，而且还可以得到，任何一个弱平稳的鞅差序列均可以写成下列的形式：

$$X_t = h_t u_t,$$

其中$\{u_t\}$是满足$E(u_1)=0$，$E(u_1^2)=1$的白噪声过程.

然而，对条件波动率建模通常还需要假设h_t是X_t的滞后序列的某种形式的函数，也按常规假设u_t在(0，1)上独立同分布，甚至假定$u_t \overset{\text{i.i.d.}}{\sim} N(0,1)$. 下面从ARCH模型开始讨论.

定义 3.7.3 设$\{u_t\}$是i.i.d.(0，1)，称$\{X_t\}$为p阶**自回归条件异方差过程**，记为ARCH(p)，$p\in\mathbf{N}$，如果

$$X_t = \sigma_t u_t,$$

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^{p} \alpha_i X_{t-i}^2,$$

134

对所有的t成立，其中$\alpha_0>0$和α_1，…，$\alpha_p\geqslant 0$均为参数.

首先详细地讨论$p=1$的情形. 设X_0为随机变量，满足$E(X_0^2)<\infty$，递推地定义序列X_t，$t\geqslant 1$如下

$$X_t = \sqrt{\alpha_0 + \alpha_1 X_{t-1}^2 u_t},\quad t = 1,2,\cdots.$$

注意到$X_1=f(X_0, u_1)$，$X_2=f(X_1, u_1)=f(f(X_0, u_1), u_1)$，以此类推，其中$f(x, y)=\sqrt{\alpha_0+\alpha_1 xy}$. 这表明$X_t$是$X_0$，$u_1$，…，$u_t$的函数. 特别地，只要$t<r$，$X_t$和$u_r$就是相互独立的. 同理，$\sigma_t$是$X_0$，$u_1$，…，$u_{t-1}$的函数. 易证

$$E(X_t|\mathcal{F}_{t-1}) = \sigma_t E(u_t) = 0,$$

这表明，只要$E|X_t|<\infty$，$\{u_t\}$就是一个鞅差序列(见后面). 进一步，σ_t^2是X_t的条件方差，因为$E(X_t^2|\mathcal{F}_{t-1})=\sigma_t^2$，再由$u_{t+h}$与$\sigma_{t+h}$，$X_t(h\geqslant 1)$相互独立知，对所有$h\geqslant 1$

$$E(X_t X_{t+h}) = E(X_t \sigma_{t+h} u_{t+h}) = E(X_t \sigma_{t+h}) E(u_t + h) = 0,$$

最后，注意到二阶矩序列$E(X_t^2)$满足下列的递推方程

$$E(X_t^2) = \alpha_0 + \alpha_1 E(X_{t-1}^2),$$

于是解得

$$E(X_t^2) = E(X_0^2) = \frac{\alpha_0}{1-\alpha_1}.$$

根据以上讨论，则有下列结论.

命题 3.7.4 对任意满足$E(X_0^2)=\dfrac{\alpha_0}{1-\alpha_1}$的初值$X_0$，ARCH(1)方程有平稳的因果解，且解是一个白噪声过程.

在实际中，当p的值较大时，ARCH(p)模型能够较好地拟合金融收益序列. 根据模型假设，σ_t^2是由过去的p个值X_{t-1}^2，…，X_{t-p}^2的加权平均给出的，这启发我们可以用更少的过去的$\sigma_t s$来替换这些项. 因此导出下列定义.

定义 3.7.5 称时间序列$\{X_t\}$为阶为(p，q)的**广义自回归条件异方差**过程，记为GARCH(p，q)，$p\geqslant 1$，$q\geqslant 0$，如果

$$X_t = \sigma_t u_t,$$

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^{p} \alpha_j X_{t-i}^2 + \sum_{j=1}^{q} \beta_j \sigma_{t-j}^2,$$

135

其中 α_0，α_1，…，$\alpha_p \geqslant 0$ 和 β_0，β_1，…，$\beta_q \geqslant 0$ 是常系数，且 α_p，$\beta_q > 0$，$u_t \overset{\text{i. i. d.}}{\sim} (0, 1)$ 独立于 $\{X_{t-k}: k \geqslant 1\}$.

实验表明：带小的 p 和 q 值的 GARCH(p, q)模型可以很好地拟合高波动率的情形，而如果使用 ARCH(p)模型的话，则需要比较大的 p.

现考虑 $p=q=1$ 的简单情形. GARCH(1, 1)模型在实际中应用会产生令人满意的模型拟合，下列定理给出了严平稳解存在的一个充分必要条件.

定理 3.7.6 设 $\alpha_0 > 0$，α_1，$\beta_1 \geqslant 0$，GARCH(1, 1)方程存在严平稳解的充分必要条件是

$$E\log(\alpha_1 u_1^2 + \beta_1) < 0.$$

此时 σ_t^2 的一个严平稳解为

$$\sigma_t^2 = \sigma_0 \Big(1 + \sum_{j=1}^{\infty} \prod_{i=1}^{j} (\alpha_1 u_{t-i}^2 + \beta_1)\Big), \quad t \in \mathbf{Z}. \tag{3.12}$$

证明 只证充分性并求出 σ_t^2 的表达式. 考虑

$$\sigma_t^2 = \alpha_0 + \alpha_1 X_{t-1}^2 + \beta_1 \sigma_{t-1}^2.$$

在上述方程中代入 $X_{t-1}^2 = \sigma_{t-1}^2 u_{t-1}^2$，则得到下列的递推式

$$\sigma_t^2 = \alpha_0 + \alpha_1 \sigma_{t-1}^2 u_{t-1}^2 + \beta_1 \sigma_{t-1}^2 = \alpha_0 + (\alpha_1 u_{t-1}^2 + \beta_1) \sigma_{t-1}^2.$$

该递推方程的一般形式为

$$x_t^2 = \alpha_0 + \gamma_{t-1} x_{t-1}^2,$$

下面证明该方程的解为级数

$$x_t^2 = \alpha_0 \Big(1 + \sum_{j=1}^{\infty} \prod_{j=1}^{i} \gamma_{t-i}\Big),$$

136

当然，前提是该级数收敛. 事实上，

$$\begin{aligned}
\alpha_0 + \gamma_{t-1} x_{t-1}^2 &= \alpha_0 + \gamma_{t-1} \alpha_0 \Big(1 + \sum_{j=1}^{\infty} \prod_{i=1}^{j} \gamma_{t-1-i}\Big) \\
&= \alpha_0 + \gamma_{t-1} \alpha_0 + \alpha_0 \gamma_{t-1} (\gamma_{t-2} + \gamma_{t-2} \gamma_{t-3} + \cdots) \\
&= \alpha_0 + \alpha_0 (\gamma_{t-1} + \gamma_{t-1} \gamma_{t-2} + \cdots) \\
&= \alpha_0 + \alpha_0 \sum_{j=1}^{\infty} \prod_{i=1}^{j} \gamma_{t-i} \\
&= x_t^2.
\end{aligned}$$

为了证明随机级数(3.12)的收敛性，考虑相关级数 $\sum_{j=1}^{\infty} \prod_{i=1}^{j} (\alpha_1 u_{t-i}^2 + \beta_1)$ 的部分和

$$\sum_{j=1}^{n} \prod_{i=1}^{j} (\alpha_1 u_{t-j}^2 + \beta_1) = \sum_{j=1}^{n} \rho^{-j} Z_j, \quad Z_j = \rho^j \prod_{i=1}^{j} (\alpha_1 u_{t-i}^2 + \beta_1),$$

其中 ρ 为任意实数. 不妨选取 $\rho>1$, 如果可以证明下列式子以概率 1 成立

$$Z_j = \rho^j \prod_{i=1}^{j} (\alpha_1 u_{t-j}^2 + \beta_1) \to 0, \quad j \to \infty, \tag{3.13}$$

则前面的部分和有有限上界 $\sum_{j=1}^{n} \rho^{-j} \to \frac{1}{1-\rho^{-1}}$, 因此得证其 a. s. 收敛. 下面验证(3.13)式, 注意到 $E\log(\alpha_1 u_1^2+\beta_1)<0$ 意味着存在某个 $\rho>1$, 使得

$$\log\rho + E\log(\alpha_1 u_1^2 + \beta_1) < 0.$$

因为 u_t 是 i. i. d. 而且 $\log(\alpha_1 u_1^2+\beta_1)$是可积的, 则当 $n\to\infty$时, 由强大数定律知 $\log(\rho) + \frac{1}{n}\sum_{i=1}^{n}\log(\alpha_1 u_{t-i}^2 + \beta_1)$(a. s.) 收敛于 $\log\rho+E\log(\alpha_1 u_1^2+\beta_1)$. 从而

$$\log\left(\rho^n \prod_{i=1}^{n} (\alpha_1 u_{t-i}^2 + \beta_1)\right) = n\left(\log(\rho) + \frac{1}{n}\sum_{i=1}^{n}\log(\alpha_1 u_{t-i}^2 + \beta_1)\right)$$

以概率 1 收敛于$-\infty$. 因此当 $n\to\infty$时, 有

$$\rho^n \prod_{i=1}^{n} (\alpha_1 u_{t-i}^2 + \beta_1) \xrightarrow{\text{a. s.}} 0,$$

137 (3.13)式得证. ■

该结论可以推广到 GARCH(p, q)模型中去, 严平稳解$\{X_t\}$的二阶矩 EX_t^2 存在, 当且仅当

$$\sum_{i=1}^{p}\alpha_i + \sum_{j=1}^{q}\beta_j < 1.$$

此时$\{X_t\}$是一个白噪声过程, 边际方差为

$$E(X_t^2) = \frac{\alpha_0}{1 - \sum_{i=1}^{p}\alpha_i - \sum_{j=1}^{q}\beta_j}.$$

ARCH 和 GARCH 模型的参数估计通常是使用极大似然估计法. 对某个均值为 0, 方差为 1 的密度函数 f, 假设扰动项序列$\{u_t\}$满足 $u_t \overset{\text{i. i. d.}}{\sim} f$. 对一个 GARCH(1, 1)模型, 在给定 X_{t-1}, …, X_0 的条件下, X_t 的条件密度函数为

$$f_{X_t|X_{t-1},\cdots,X_0}(x_t|x_{t-1},\cdots,x_0) = f_{X_t|X_{t-1}}(x_t \mid x_{t-1}) = \frac{1}{\sigma_t}f\left(\frac{x_t}{\sigma_t}\right),$$

因此, 似然函数为

$$L(\alpha_0,\alpha_1,\beta_1) = \prod_{t=1}^{T}\frac{1}{\sigma_t}f\left(\frac{x_t}{\sigma_t}\right),$$

其中 $\sigma_t = \sqrt{\alpha_0 + \alpha_1 X_{t-1}^2 + \beta_1\sigma_{t-1}^2}$可由递推方法计算. 注意到我们需要一个已知的初值 σ_0, 通常是使用历史数据的样本标准差代替, 或简单地设 $\sigma_0=0$. 对一个 GARCH(p, q)模型, 其似然函数可以类似地得到, 但是此时需要给定初值 X_{-p+1}, …, X_0 和 σ_{-p+1}, …, σ_0.

得到了 GARCH 模型未知参数的估计值, 就可以得到估计量$\hat{\sigma}_t$, 例如对 GARCH(1, 1)模型, 有$\hat{\sigma}_t = \sqrt{\hat{\alpha}_0 + \hat{\alpha}_1 X_{t-1}^2 + \hat{\sigma}_{t-1}}$. 从而可以得到下列残差估计量

$$\hat{u}_t = \frac{\hat{X}_t}{\hat{\sigma}_t},$$

如果模型是正确的，该残差是一个近以白噪声过程，这可以通过残差的样本自协方差来验证.

经典 GARCH(p, q)模型的修正和推广已经被很多文献讨论过了，下面给出三个例子. 考虑到不对称信息对波动率的影响，已有很多论文讨论用$(X_{t-1}+\delta|X_{t-1}|)^2$ 来代替模型中的 X_{t-1}^2项，其中参数 $\delta\in[0, 1]$. 因此，如果 $X_{t-1}\geqslant 0$，σ_t^2 表达式右边的 X_{t-1}^2变为$(1+\delta)^2X_{t-1}^2$，否则为$(1-\delta)^2X_{t-1}^2$. 该杠杆效应的经济学解释是：当公司的股票下跌时，它的资产负债率会上升，从而引起波动率上升，因为波动率是反映股票投资风险的指标.

138

GARCH 均值模型是下列形式的模型：

$$X_t = \beta Z_t + \lambda\sigma_t + \sigma_t u_t,$$

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^{p}\alpha_i X_{t-i}^2 + \sum_{j=1}^{q}\beta_j\sigma_{t-j}^2,$$

其中 Z_t 是一个回归自变量. 显然 X_t 关于$\mathcal{F}_{t-1}$的条件均值为 $\beta Z_t+\lambda\sigma_t$，它依赖于波动率.

指数 GARCH 模型假定条件方差为

$$\sigma_t^2 = \exp\Big(\alpha_0 + \sum_{j=1}^{p}\alpha_i g(X_{t-i}) + \sum_{j=1}^{q}\beta_j\log(\sigma_{t-j})\Big),$$

其中 α_0，α_1，…，α_p 和 β_1，…，β_q 均是实值参数. 如果用对数波动率建模，参数不再具有这样形式. 式中的 $g(X_{t-i})$为

$$g(X_t) = \theta X_t + \gamma(|X_t| - E|X_t|),$$

其中 θ 和 γ 是新的参数. 可以得到下列 MA(∞)表达式

$$\log(\sigma_t^2) = \omega_t + \sum_{k=1}^{\infty} g(X_{t-k}).$$

如果 $\sum_{k=1}^{\infty}\beta_k^2 < \infty$，则 σ_t^2 是严平稳的.

3.8 长记忆序列

研究表明，一些金融时间序列具有长记忆效应，即它们的自相关系数比 AR 和 ARMA 模型衰减得更慢. 已经知道，自协方差 γ_k 以指数速率 $\gamma_k \sim a^{-k}$衰减为 0，当 k 变大时迅速衰减. 一般说这种模型具有短记忆效应. 但是，如果 $\gamma_k\sim k^{-\beta}$，$\beta>0$，则 γ_k 衰减要慢得多. 如 $\beta>1$，它们还是可求和的，但对 $0<\beta<1$ 连这种性质也没有了.

为了便于理解长记忆过程，首先介绍分数阶差分.

3.8.1 分数阶差分

对 $d\in\mathbf{N}$，d 阶差分为滞后多项式$(1-L)^d$ 作用的结果，即

$$(1-L)X_t = X_t - X_{t-1},$$

139

$$(1-L)^2 X_t = X_t - 2X_{t-1} + X_{t-2},$$

如此等等．如果$-1<d<1$，则称为**分数阶差分**．

对 $d\in\mathbf{R}$，滞后算子 $p(L)=(1-L)^d$ 的级数表达式可以通过函数 $f(z)=(1-z)^d$，$z\in\mathbf{C}$ 的泰勒级数展开得到．因为

$$f^{(k)}(z) = \left(\prod_{i=0}^{k-1}(d-i)\right)(-1)^k(1-z)^{d-k}, \quad k\in\mathbf{N}$$

因此有下列二项级数，

$$(1-z)^d = \sum_{k=0}^{\infty}\psi_k z^k, \quad z\in\mathbf{C}, \tag{3.14}$$

其中系数 $\psi_0=1$，且有

$$\psi_k = \frac{f^{(k)}(0)}{k!} = \frac{(-1)^k}{k!}\prod_{i=0}^{k-1}(d-i) = (-1)^k\binom{d}{k}, \quad k\in\mathbf{N},$$

其中，

$$\binom{d}{k} = \frac{d(d-1)\cdots(d-k+1)}{k!}$$

为广义二项系数．对 $d>0$，级数在单位圆盘上绝对收敛，因而算子

$$(1-L)^d = \sum_{k=0}^{\infty}\psi_k L^k$$

有定义，是一个 L_1 滤子．因此，可以考虑形如$(1-L)^d\varepsilon_t$ 的线性过程，其中 ε_t 是白噪声过程．更一般地，如果过程$\{X_t\}$存在，且 $\sup_n E|X_n|<\infty$，且 $\sup_n E|X_n|^2<\infty$（$\{X_t\}$属于 L_2），则过程$(1-L)^d X_t$ a. s. 存在．

在考虑$-1<d<0$ 的情形之前，先讨论系数 ψ_k 的一些性质．

系数 ψ_k 可以表示成 Γ 函数的形式．为了证明这一点，先记

$$\psi_k = \frac{1}{k!}\prod_{i=0}^{k-1}(-d+i). \tag{3.15}$$

回顾一下 Γ 函数的定义，

$$\Gamma(z) = \lim_{n\to\infty}\frac{n^z n!}{z(z+1)(z+2)\cdots(z+n)}, \quad z\in\mathbf{C}\setminus\{0,-1,-2,\cdots\}.$$

对于正实数 z，有下列积分表达式

140

$$\Gamma(z) = \int_0^{\infty}\exp(-t)t^{z-1}\mathrm{d}t.$$

Γ 函数有下列性质：$\Gamma(n+1)=n!$，$n\in\mathbf{N}$，更一般地，在其定义域上有递推式 $\Gamma(z+1)=z\Gamma(z)$．因此可以得到下列公式

$$\Gamma(z+n+1) = (z+n)\Gamma(z+n) = \cdots = \prod_{i=0}^{n}(z+i)\Gamma(z),$$

利用上式可将(3.15)的乘积写成 Γ 函数的形式．事实上，设 $z=-d$，由 $1/k!=1/\Gamma(k+1)$得到

$$\psi_k=\frac{\Gamma(k-d)}{\Gamma(k+1)\Gamma(-d)},\quad k\in\mathbf{N}.$$

下面的引理表明，系数 ψ_k 以速率 $k^{-\gamma}$ 衰减，其中 $\gamma=d+1>0$．证明要用到 Sterling 公式：$\Gamma(x)\sim\sqrt{2\pi}\mathrm{e}^{-x+1}(x-1)^{x-1/2}$，$x\to\infty$.

引理 3.8.1 对 $d>-1$，当 $k\to\infty$ 时，有 $\psi_k\sim\dfrac{k^{-(d+1)}}{\Gamma(-d)}$.

证明 由 Sterling 公式，有

$$\begin{aligned}\psi_k&=\frac{\Gamma(k-d)}{\Gamma(k+1)\Gamma(-d)}\\&\sim\frac{\mathrm{e}^{d-k+1}(k-d-1)^{k-d-1/2}}{\Gamma(-d)\mathrm{e}^{-k}k^k+1/2}\\&=\frac{1}{\Gamma(-d)}\mathrm{e}^{d+1}\left(\frac{k-d-1}{k}\right)^k\frac{(k-d-1)^{-d-1/2}}{k^{1/2}},\end{aligned}$$

其中用到当 $k\to\infty$时，

$$\left(\frac{k-d-1}{k}\right)^k=\left(1-\frac{d+1}{k}\right)^k\to\mathrm{e}^{-(d+1)}.$$

又因为 d 是固定的值，则当 $k\to\infty$时，

$$\begin{aligned}\frac{(k-d-1)^{-d-1/2}}{k^{1/2}}&=(k-d-1)^{-(d+1)}\left(\frac{k-d-1}{k}\right)^{1/2}\\&=k^{-(d+1)}\left(\frac{k-(d+1)}{k}\right)^{-(d+1)}\left(\frac{k-(d+1)}{k}\right)^{1/2}\\&\sim k^{-(d+1)},\end{aligned}$$

证毕． ■

为了研究 $1-L$ 的负分数次方，需要下列引理．

141

引理 3.8.2 对 $-1/2<d<1/2$，有 $\sum\limits_{k=0}^{\infty}\psi_k^2<\infty$，使得 $\sum\limits_{k=0}^{\infty}\psi_kL^k$ 是一个 L_2 滤波算子．

证明 设 $\varepsilon>0$，由引理 3.8.1 知，存在 k_0 使得 $|\psi_k/(k^{-(d+1)}/\Gamma(-d))|\leqslant 1+\varepsilon$ 对所有的 $k\geqslant k_0$ 成立．对某个常数 $0<C_1<\infty$，有

$$\begin{aligned}\sum_{k=0}^{\infty}\psi_k{}^2&=\sum_{k=0}^{k_0-1}\psi_k{}^2+\sum_{k=k_0}^{\infty}\left(\frac{1}{\Gamma(-d)}\frac{1}{k^{d+1}}\right)^2\left(\frac{\psi_k}{k^{-(d+1)}/\Gamma(-d)}\right)^2\\&\leqslant C_1+(1+\varepsilon)^2\sum_{k=k_0}^{\infty}\left(\frac{1}{\Gamma(-d)}\frac{1}{k^{d+1}}\right)^2.\end{aligned}$$

余下来只需证明 $\sum\limits_{k=k_0}^{\infty}\left(\frac{1}{k^{d+1}}\right)^2<\infty$ 即可．若 $0\leqslant d<1/2$，则

$$\sum_{k=k_0}^{\infty}\left(\frac{1}{k^{d+1}}\right)^2\leqslant\sum_{k=0}^{\infty}\frac{1}{k^{2+2d}}\leqslant\sum_{k=0}^{\infty}\frac{1}{k^2}<\infty.$$

若$-1/2<d<0$，取 d_0 使得$-1/2<d_0<d$，于是 $2d_0\in(-1,0)$，$\delta=1+2d_0>0$. 从而有

$$\sum_{k=k_0}^{\infty}\frac{1}{k^{2+2d}}\leqslant\sum_{k=k_0}^{\infty}\frac{1}{k^{2+2d_0}}\leqslant\sum_{k=0}^{\infty}\frac{1}{k^{1+\delta}}<\infty,$$

证毕. ■

已经定义了分数阶差分算子$(1-L)^d$，$d>0$，将其作用于白噪声过程，则得到平稳时间序列. 现考虑它的逆算子$(1-L)^{-d}$. 对大于-1的负指数，二项级数仅当$|z|<1$时收敛. 考虑方程(3.6)，即 $\psi(z)\theta(z)=1$，其中$|z|<1$. 对 $0<d<1$，其解 $\theta(z)=(1-z)^{-d}$是一个解析函数. 形如(3.14)的 Taylor 展式为

$$\theta(z)=\sum_{k=0}^{\infty}\theta_k z^k,\quad \theta_k=\theta^{(k)}(0)/k!,$$

用$-d$代替d即得到系数为

$$\theta_k=\frac{1}{k!}\prod_{i=0}^{k-1}(d+i)=\frac{\Gamma(k+d)}{\Gamma(k+1)\Gamma(d)}.$$

如果 $0<d<1/2$，使用 Sterling 公式可以进一步得到

142

$$\theta_k\sim k^{d-1}/\Gamma(d),\quad k\to\infty,$$

这同样表明 $\sum_{k=0}^{\infty}\theta_k^2<\infty$，即 $\theta(L)$是一个 L_2 滤波. 则当 $0<d<1/2$ 时，根据命题 3.4.3，

$$X_t=(1-L)^{-d}\varepsilon_t=\sum_{k=0}^{\infty}\theta_k\varepsilon_{t-k},\quad \varepsilon_t\sim \mathrm{WN}(0,\sigma^2),\quad \sigma^2\in(0,\infty),$$

在 L_2 意义下存在，从而可以得出方程$(1-L)^dX_t=\varepsilon_t$ 对应的平稳解.

定义 3.8.3(分整噪声) 设$-1/2<d<1/2$，$\{\varepsilon_t\}$为白噪声过程，则方程$(1-L)^dX_t=\varepsilon_t$的平稳解过程$\{X_t\}$称为**分整噪声**.

在引理 3.5.2 里已经看到，由滞后算子 $p(L)$给出的线性过程的谱密度为 $f(\omega)=\frac{\sigma^2}{2\pi}|p(\mathrm{e}^{\mathrm{i}\omega})|^2$，有了谱密度，就可以计算出样本自协方差函数. 根据定理 3.5.7，计算公式为 $\gamma(h)=\int_{-\pi}^{\pi}\cos(\omega h)f(\omega)\mathrm{d}\omega$.

对 $0<d<1/2$，线性滤波 $\psi(L)=\sum_j\psi_jL^j$ 是一个 L_1 滤波. 由 $\psi(z)=(1-z)^{-d}$，可以得到序列 $X_t=(1-L)^{-d}\varepsilon_t=\psi(L)\varepsilon_t$ 的谱密度为

$$f_X(\omega)=\frac{\sigma^2}{2\pi}|1-\mathrm{e}^{\mathrm{i}\omega}|^{-2d},$$

利用等式$|1-\mathrm{e}^{\mathrm{i}\omega}|=2\sin(\omega/2)$，进一步地有

$$f_X(\omega)=\frac{\sigma^2}{2\pi}|1-\sin(\omega/2)|^{-2d}.$$

可以利用上述的公式来计算自协方差:

$$\gamma_X(h)=\int_{-\pi}^{\pi}\mathrm{e}^{\mathrm{i}h\omega}f(\omega)\mathrm{d}\omega$$

$$= \frac{\sigma^2}{\pi}\int_0^{\pi}\cos(h\omega)(2\sin(\omega/2))^{-2d}\mathrm{d}\omega.$$

其中最后一个积分具有$\int_0^{\pi}\cos(hx)\sin^{z-1}(x)\mathrm{d}x$的形式，其结果为$\frac{\pi\cos(h\pi/2)\Gamma(z+1)2^{1-z}}{z\Gamma((z+h+1)/2)\Gamma((z-h-1)/2)}$.

引理 3.8.4 设$0<d<1/2$，$\varepsilon_t \sim \mathrm{WN}(0, \sigma^2)$，则序列$\{X_t\}$的样本自协方差和自相关函数分别为

$$\gamma_X(h) = \sigma^2 \frac{(-1)^h\Gamma(1-2d)}{\Gamma(h-d+1)\Gamma(1-h-d)},$$

143

$$\rho_X(h) = \frac{\Gamma(h+d)\Gamma(1-d)}{\Gamma(h-d+1)\Gamma(d)},$$

这里$h\in\mathbf{Z}$. 而且进一步地，有

$$\rho_X(h) \sim h^{2d-1}\frac{\Gamma(1-d)}{\Gamma(d)}.$$

长记忆过程的另一种引入方式为：均值为零的高斯序列$\{Z_t\}$是长相依的，如果它的样本自协方差函数$\gamma_Z(k)$满足

$$\gamma_Z(k) \sim k^{-D}L(k), \quad k\to\infty,$$

其中$0<D<1$为某个定常数，$L(x)$为一个缓变函数，即对任意$c>0$，均有$\lim_{x\to\infty}\frac{L(cx)}{L(x)}=1$. 此时也称$H=1-D/2$为 **Hurst 指数**. 引理 3.8.4 表明，满足$0<d<1/2$的分整序列的 Hurst 指数为$H=1/2+d$.

3.8.2 分整过程

设$\{X_t\}$为一个时间序列，如果它的分数阶差分可以表示为一个 ARMA(p, q)过程，则称其为一个 FARIMA(p, q)过程. 其严格定义如下.

定义 3.8.5(FARIMA 过程) 时间序列$\{X_t\}$称为(p, d, q)阶分整 ARMA 过程，其中$d\in(-1/2, 1/2)$，如果$\{X_t\}$是下列方程的平稳解：

$$\psi(L)(1-L)^{d}X_t = \phi(L)\varepsilon_t,$$

其中$\{\varepsilon_t\}$是一个白噪声；ψ，ϕ分别为p，q阶滞后多项式.

3.9 评注与延伸阅读

关于离散时间的鞅，可阅读 Williams(1991). 有关鞅及其极限理论比较经典的著作是 Hall 和 Heyde(1980). Brockwell 和 Davis(1991)介绍了时间序列分析及模型参数估计. 要深入了解金融市场理论，可参考 Fan 和 Yao(2003)，Carmona(2004)，Lai 和 Xing(2008)，以及 Tsay(2010)和 Jondeau 等(2007). ARCH 模型的启发性工作始于 Engle(1982)，指数 GARCH 模型源于 Nelson(1991). 对这些模型及它们的参数估计进行综合性论述的是 Straumann(2005). 定理 3.7.6 的证明是参考 Kreiss 和 Neuhaus(1990)，但其主要思想源于 Nelson(1990).

144

参考文献

Brockwell P.J. and Davis R.A. (1991) *Time Series: Theory and Methods*. Springer Series in Statistics 2nd edn. Springer-Verlag, New York.

Carmona R.A. (2004) *Statistical Analysis of Financial Data in S-Plus*. Springer Texts in Statistics. Springer-Verlag, New York.

Engle R.F. (1982) Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica* **50**(4), 987–1007.

Fan J. and Yao Q. (2003) *Nonlinear Time Series: Nonparametric and Parametric Methods*. Springer Series in Statistics. Springer-Verlag, New York.

Hall P. and Heyde C.C. (1980) *Martingale Limit Theory and its Application*. Academic Press Inc. [Harcourt Brace Jovanovich Publishers], New York. Probability and Mathematical Statistics.

Jondeau E., Poon S.H. and Rockinger M. (2007) *Financial Modeling under Non-Gaussian Distributions*. Springer Finance. Springer-Verlag London Ltd., London.

Kreiss J.P. and Neuhaus G. (2006) *Einführung in die Zeitreihenanalyse*. Springer, Berlin Heidelberg.

Lai T.L. and Xing H. (2008) *Statistical Models and Methods for Financial Markets*. Springer Texts in Statistics. Springer, New York.

Nelson D.B. (1990) Stationarity and persistence in the GARCH(1, 1) model. *Econometric Theory* **6**(3), 318–334.

Nelson D.B. (1991) Conditional heteroskedasticity in asset returns: a new approach. *Econometrica* **59**(2), 347–370.

Straumann D. (2005) *Estimation in conditionally heteroscedastic time series models* vol. 181 of *Lecture Notes in Statistics*. Springer-Verlag, Berlin.

Tsay R.S. (2010) *Analysis of Financial Time Series*. Wiley Series in Probability and Statistics 3rd edn. John Wiley & Sons Inc., Hoboken, NJ.

145

Williams D. (1991) *Probability with Martingales*. Cambridge Mathematical Textbooks. Cambridge University Press, Cambridge.

第 4 章　多期模型的套利理论

第 2 章已经讨论了离散时间的鞅理论，本章讨论未定权益在离散的固定时刻进行交易的条件下，未定权益的定价问题. 不失一般性，不妨设离散时刻为 $t=0, \cdots, T$，其中 T 既表示交易期数，又表示终止时刻. 在介绍无套利与存在等价鞅测度 P^* 的关系以前，首先必须将单期模型的一些概念，如**自融资策略**、**等价鞅测度**和**无套利**推广到多期模型. 在本章将会看到，如果金融市场的样本空间 Ω 为有限集，则可以给出 P^* 漂亮的计算公式.

在讨论了多期情形的一般理论之后，本章重点讨论 Cox-Ross-Rubinstein 二叉树模型，在该模型中可以求出所有变量的显性表达式，包括路径独立衍生产品的 delta 对冲. 通过构造一列渐近等价鞅测度，研究股价在这些测度下的收敛分布将会发现，多期二叉树模型是导出著名 Black-Scholes 期权定价公式的简单而有力的工具.

最后，本章介绍如何定价美式衍生品，比如股票美式期权，它是一种持有者可以在到期日之前的任何时刻交易的衍生品. 利用第 3 章的最优停时理论，本章用一种简单易行的方式处理美式期权的定价问题，得到的结果是非常漂亮的. 利用二叉树理论本章还导出了美式衍生品无套利定价的基本算法.

4.1　定义与预备

在本章中，将金融市场的模型表示如下：设$(\Omega, \mathcal{F}, \{\mathcal{F}_t\}, P)$为带滤概率空间，其中满足

$$(\varnothing, \Omega) = \mathcal{F}_0 \subset \mathcal{F}_1 \subset \cdots \subset \mathcal{F}_T \subset \mathcal{F}$$

P 为实概率测度. 设投资者可以投资 d 项不同的风险资产，也可选择将资金存入银行账户，则价格过程为下列 F_t-适应过程：

$$\{S_{ti} : t = 0, 1, \cdots, T\}, \quad i = 0, \cdots, d,$$

其中，$0 \leqslant S_{ti}$ 和 $E(S_{ti}) < \infty$ 对所有的 t 和 i 成立. 与第 2 章类似，假定每期存款或贷款利率是相同的固定值. 为了简化讨论，设利率为已知且固定的常数，即，t 时刻存入的款项 C 在 $t+1$ 时刻将增长为 $C(1+r)$，r 就是每期的利率，并且 r 对不同的 t 保持不变. 因此，银行账户可以等价地用债券来描述，其价格过程为

$$S_{t0} = S_{00}(1+r)^t, \quad t = 0, \cdots, T,$$

其中，S_{00} 是债券的初始面值.

需要指出的是，尽管为了简化讨论对债券模型作了上述限制，但是本章的方法对所谓的局部无风险债券模型同样适用.

定义 4.1.1　称价格过程为

$$S_{t0} = \prod_{i=1}^{t}(1+r_i), \quad t = 0, \cdots, T,$$

的债券模型为**局部无风险债券模型**，其中 $\{r_t\}$ 为可预报过程.

4.2 自融资交易策略

因为投资者可以交易的时刻为 0，$\cdots T-1$，一个投资策略由 t 时刻到 $t+1$ 时刻持有每种资产 i 的份额 φ_{ti} 确定，显然 φ_{Ti} 是终止时刻的资产份额. 也就是说，投资者在 0 时刻确定持有每种资产 i 的头寸. 如果 $\varphi_{1i}S_{1i}>0$，表示他支付了 $\varphi_{1i}S_{1i}$，持有资产 i 的多头头寸；反之若 $\varphi_{1i}S_{1i}<0$，表示他收到金额 $|\varphi_{1i}S_{1i}|$，持有资产 i 的空头寸并. 换句话说，建立了投资组合 $\varphi_1=(\varphi_{10},\cdots,\varphi_{1d})'$，投资者在下一时刻拥有或建立的资产组合为 $\varphi_2=(\varphi_{20},\cdots,\varphi_{2d})'$，以此类推.

因为 φ_{ti} 是由直到 $t-1$ 时刻为止所包含的所有信息决定的，下列定义的引入是自然的.

定义 4.2.1 一个满足 $\varphi_0=\varphi_1$，取值于 $\mathbf{R}^{d+1}$ 的可预报过程 $\{\varphi_t:t=0,\cdots,T\}$，$\varphi_t=$ 148 $(\varphi_{t0},\cdots,\varphi_{td})'$ 称为**交易策略**.

注意，令 $\varphi_0=\varphi_1$ 主要是为了简化结果. 在以下的模型中，φ_0 实际上是不需要的.

定义 4.2.2 令 $\{S_t:t=0,\cdots,T\}$ 为一个 d 维价格过程，$\varphi=\{\varphi_t:t=0,\cdots,T\}$ 为交易策略，则由

$$V_t=\varphi_t'S_t,\quad t=0,\cdots T,$$

确定的过程 $V_t=V_t(\varphi)$ 称为交易策略 φ_t 的**价值过程**.

V_0 是交易策略运行所需的初始资金，$V_1=\varphi_1'S_1$ 为交易策略在 1 时刻的价值，一般地，$V_t=\varphi_t'S_t$ 为交易策略在 t 时刻的价值. 在现实操作中，投资策略可能会涉及额外的资金流入或流出，比如向投资者支付红利或新的投资者注入资金等. 这些问题可以通过引入一项新的资产来解决，在本章中暂不考虑这些问题. 因此，仅考虑初始时刻融资 $\varphi_0'S_0$，后续的资产变化由投资组合自身产生，增持某项资产意味着要出售其他资产或从银行账户转账.

对任意 $t\in\{1,\cdots,T-1\}$，从 $t-1$ 至 t 持有的资产组合 φ_t 的价值 $\varphi_t'S_t$ 称为 t 时刻的值. 在 $t-1$ 时刻建仓需要本金 φ'_tS_{t-1}，在 t 时刻清仓可以实现价值 φ'_tS_t. 在下一期，需要确定策略 φ_{t+1}，所需成本为 $\varphi_{t+1}'S_t$，则净增值为

$$C_t=\varphi_{t+1}'S_t-\varphi_t'S_t=(\varphi_{t+1}-\varphi_t)'S_t=\Delta\varphi_{t+1}'S_t.$$

其中，差分算子 Δ 定义为

$$\Delta\varphi_{t+1}=\varphi_{t+1}-\varphi_t.$$

如果 $C_t>0$，则需要额外的融资来构建资产组合 φ_{t+1}；反之，若 $C_t<0$，则表示从组合中流出资金 $|C_t|$. 当 $C_t=0$ 时，称投资策略为自融资策略.

定义 4.2.3 一个交易策略称为**自融资的**，如果对所有的 $t=1,\cdots,T$ 均有

$$\Delta\varphi_t'S_{t-1}=0\Leftrightarrow\varphi_t'S_{t-1}=\varphi_{t-1}'S_{t-1}$$

由离散时间的随机积分的定义，设 φ_t 和 S_t 均为取值于 $\mathbf{R}$ 的过程，φ_t 是可预报的，S_t 是适应的，则离散时间的随机积分 $\int\varphi_r\mathrm{d}S_r$ 为下列 $\mathbf{R}$ 值过程

149

$$I_t=\int_0^t\varphi_r\mathrm{d}S_r=\sum_{r=1}^{t}\varphi_r\Delta S_r=\sum_{r=1}^{t}\varphi_r(S_r-S_{r-1}),\quad t=0,\cdots,T,\tag{4.1}$$

(见定义 3.2.11). 如果 φ_t 和 S_t 均为 k 维的, 则 $\int \varphi_r' \mathrm{d}S_r$ 为下列 $\mathbf{R}$ 值过程

$$I_t = \int_0^t \varphi_r' \mathrm{d}S_r = \sum_{r=1}^{t} \varphi_r'(S_r - S_{r-1}), \quad t = 0, \cdots, T.$$

定理 4.2.4 交易策略 $\{\varphi_t : t=0, \cdots, T\}$ 是自融资的, 当且仅当对所有的 $t=0, \cdots, T$, 均有

$$V_t = V_0 + \sum_{r=1}^{t} \varphi_r' \Delta S_r = V_0 + \int_0^t \varphi_r' \mathrm{d}S_r$$

其中, $V_0 = \varphi_0{}' S_0$. 换句话说: 一个交易策略是自融资的等价于对应的价值过程是一个离散时间的 Itô随机积分.

证明 当 $t=0$ 时结论显然成立. 又因为

$$\begin{aligned} \varphi_t \text{ 是自融资的} &\Leftrightarrow \Delta\varphi_t' S_{t-1} = 0, \quad t = 1, \cdots, T \\ &\Leftrightarrow \varphi_t' S_{t-1} - \varphi_{t-1}' S_{t-1} = 0, \quad t = 1, \cdots, T. \end{aligned}$$

上式两边同时加 $\varphi_t' S_t - \varphi_t' S_{t-1}$, 得

$$\varphi_t' S_t - \varphi_{t-1}' S_{t-1} = \varphi_t' S_t - \varphi_t' S_{t-1} = \varphi_t'(S_t - S_{t-1})$$

上式左边是 $V_t = V_0 + \sum_{r=1}^{t} \Delta V_r$ 的第 $t+1$ 个被加数, 即 ΔV_t. 因此, 对 $t=1, \cdots, T$, 有

$$V_t = V_0 + \sum_{r=1}^{t} \Delta V_r = V_0 + \int_0^t \varphi_r' \mathrm{d}S_r. \qquad \blacksquare$$

定义 4.2.5 称过程 $S_t^* = (1, S_{t1}/S_{t0}, \cdots, S_{td}/S_{t0})'$ $(t=0, \cdots, T)$ 为**贴现价格过程**, 称 $V_t^* = \varphi_t' S_t^*$ $(t=0, \cdots, T)$ 为投资策略的**贴现价值过程**.

注 4.2.6 若银行账户(标准计价单位)为初始面值为 1, 固定利率为 r 的债券, 则有

$$S_{ti}^* = \frac{S_{ti}}{(1+r)^t}, \quad V_t^* = \frac{\varphi_t' S_t}{(1+r)^t}.$$

注意到 $\Delta\varphi_t' S_{t-1} = \varphi_t' S_{t-1} - \varphi_{t-1}' S_{t-1} = 0$ 等价于 $\Delta\varphi_t' S_{t-1}^* = 0$, 所以有下列命题.

命题 4.2.7 交易策略是自融资的当且仅当其对应的价值过程 $V_t^* = V_t^*(\varphi)$ 对 $t=0, \cdots, T$, 有

$$V_t^* = V_0^* + \sum_{r=1}^{t} \varphi_r' \Delta S_r^* = V_0^* + \int_0^t \varphi_r' \mathrm{d}S_r^*.$$

150

值得注意的是, 在贴现时, 随机积分不依赖于存入银行账户的资金 φ_{t0} $(t=1, \cdots, T)$, 因为对所有的 t 均有 $\Delta S_{t0}^* = 0$.

定理 4.2.8 设 Q 为可测空间 $(\Omega, \mathcal{F})$ 上的概率测度, $\{S_{ti}^* : t=0, \cdots T\}$ $(i=1, \cdots, d)$ 是测度 Q 下的鞅, 如果 $\{\varphi_t\}$ 为自融资策略, 其对应的价值过程 $V_t = \varphi_t' S_t$ 对所有的 t 是 Q-可积的, 那么 V_t^* 是 Q 鞅. 即当 $s \leqslant t$ 时, 有

$$E_Q(V_t^* \mid \mathcal{F}_s) = V_s^* \quad Q\text{-a. s.},$$

且 $E_Q(V_t^*) = V_0$, $V_t^* = E_Q(V_T^* \mid \mathcal{F}_t)$, Q-a. s. 对 $t=1, \cdots, T$ 均成立.

证明 因为 φ_t 是自融资的, $V_{t-1} = \varphi_t' S_{t-1}$, 所以

$$\varphi_{t+1}'(S_{t+1} - S_t) = V_{t+1} - \varphi_{t+1}' S_t = V_{t+1} - V_t$$

是 Q-可积的. 由 S_t^* 的鞅性质，有

$$E_Q(\varphi_{t+1}'(S_{t+1}^* - S_t^*) | \mathcal{F}_t) = \varphi_{t+1}'(E_Q(S_{t+1}^* | \mathcal{F}_t) - S_t^*) = 0.$$

因此，对 $t=0$，…，$T-1$，有

$$\begin{aligned} E_Q(V_{t+1}^* | \mathcal{F}_t) &= E_Q\left(V_0^* + \sum_{i=1}^t \varphi_i' \Delta S_i^* \,|\, \mathcal{F}_t\right) + E(\varphi_{t+1}'(S_{t+1}^* - S_t^*) | \mathcal{F}_t) \\ &= V_0^* + \sum_{i=1}^t \varphi_i' \Delta S_i^* \\ &= V_t^*, \end{aligned}$$

又因为 φ_t 是可预报过程，即 φ_t 是$\mathcal{F}_{t-1}$可测的，所以$\{V_t^*\}$是 Q-鞅. ■

注 4.2.9 下面是 $E|V_t|<\infty$，$t=0$，…，T 的一些充分条件.

(ⅰ) Ω 是有限的.

(ⅱ) φ_t 有界. 即对 $t=0$，…，T，均有 $|\varphi_t|\leqslant C$.

(ⅲ) 对 p，$q\in[1, \infty]$且 $p^{-1}+q^{-1}=1$，均有 $E|\varphi_t|^p<\infty$，$ES_t^q<\infty$，然后由 Hölder 不等式知这是一个充分条件.

定理 4.2.8 给出了一个重要的启示：如果找到一个测度 Q，使得价格过程是 Q-鞅，那么自融资策略的贴现价值过程 V_t^* 可由终止时刻价值的贴现 V_T^* 来表出，即

151

$$V_t^* = E_Q(V_T^* | \mathcal{F}_t).$$

且 t 时刻的价值为

$$V_t = E_Q(V_T^*(1+r)^t | \mathcal{F}_t) = E_Q(V_T(1+r)^{(t-T)} | \mathcal{F}_t).$$

若自融资策略被用于复制一个未定权益，则上述公式给出了未定权益的公平价值. 下面将上述结果用于无套利定价中去.

4.3 无套利与鞅测度

定义 4.3.1 称自融资交易策略 $\varphi=\{\varphi_t\}$为**套利机会**或**套利**，如果对应的价值过程 $V_t=V_t(\varphi)$满足

$$V_0 \leqslant 0, \quad V_T \geqslant 0(P\text{-a. s.}), \quad P(V_T > 0) > 0. \tag{4.2}$$

注 4.3.2

(ⅰ) φ_t 是否存在套利机会依赖于概率测度 P. 对与 P 等价的所有概率测度，φ_t 的套利性都是一致的，且从套利机会中赚取的利润也可能与 P 有关.

(ⅱ) 显然，方程(4.2)等价于

$$V_0^* \leqslant 0, \quad V_T^* \geqslant 0(P\text{-a. s.}), \quad P(V_T^* > 0) > 0.$$

多期模型是一条单期模型链，可以建立多期套利和单期套利之间的关系.

定理 4.3.3 多期模型存在套利机会当且仅当在对应的单期模型链中至少有一个存在套利机会.

定义 4.3.4

(ⅰ) 称概率测度 Q 为**鞅测度**，如果 Q 关于$\mathcal{F}_0$ 是平凡的，即对任意的 $A\in\mathcal{F}_0$，$Q(A)$

$\in\{0,1\}$，且所有的贴现价值过程$\{S_{ti}^*:t=0,\cdots,T\}$，$i=1,\cdots,d$ 在 Q 下是$\mathcal{F}_t$ 鞅.

（ⅱ）称 Q 为**等价鞅测度**(EMM)，如果 Q 是一个鞅测度，且 $Q\sim P$.

有了上述准备，现可以得到一些基本定理，这里只对有限概率空间的情形予以证明. 按照前面的记法，有限概率空间记为 $\Omega=\{\omega_1,\cdots,\omega_n\}$，$n\in\mathbf{N}$，随机变量 X 可以写成向量形式$(X(\omega_1),\cdots,X(\omega_n))$，现将单期模型的一些结果推广到多期模型.

152

定理 4.3.5 金融市场无套利的充分必要条件是存在等价鞅测度.

证明 设 $\Omega=\{\omega_1,\cdots,\omega_n\}$，$n\in\mathbf{N}$，且 $p_i:=P(\{\omega_i\})>0$ 对 $i=1,\cdots,n$ 均成立，将随机变量记成一个 n 维向量. 将定理 2.5.4 的证明思想用于多期情形，考虑集合

$$U=\{(-V_0^*(\varphi),V_T^*(\varphi)):\varphi=\{\varphi_t\}\text{ 可测}\}\subset\mathbf{R}^{n+1},$$

其中，变量 $V_T^*(\varphi)$可以记为$(V_T^*(\varphi)(\omega_1),\cdots,V_T^*(\varphi)(\omega_n))$，$V_0^*(\varphi)$是常量. 注意到 U 是 $\mathbf{R}^{n+1}$的线性子空间. 令

$$M=\{(y_0,\cdots,y_n)'\in\mathbf{R}^{n+1}:y_i\geqslant 0,i=0,\cdots,n,\text{存在某个 } j,\text{有 } y_j>0\}.$$

则 $0\notin M$，且市场无套利当且仅当 $U\cap M=\varnothing$. 进一步，考虑

$$K=\{y\in M:y'\mathbf{1}=1\},\quad \mathbf{1}=(1,\cdots,1)'\in\mathbf{R}^{n+1},$$

则集合 K 是 M 的一个非空凸紧子集，且 $U\cap K=\varnothing$. 由定理 2.4.5 的凸集分离理论知：存在向量 $\lambda=(\lambda_0,\cdots,\lambda_n)'\in\mathbf{R}^{n+1}$，使得

$$\begin{aligned}\lambda'x=0,&\quad\forall x\in U,\\ \lambda'x>0,&\quad\forall x\in K.\end{aligned}$$

又因为任一单位向量 e_j 包含于 K，所以 $\lambda_j=\lambda'e_j>0$，$j=0,\cdots,n$. 定义概率测度

$$P^*(\{\omega_j\}):=\frac{\lambda_j}{\sum\limits_{k=1}^n\lambda_k},\quad j=1,\cdots,n,$$

则可证 P^* 等价于 P. 从而，P^* 就是要找的鞅测度.

事实上，由

$$\begin{aligned}E^*|S_{ti}|&=\sum_{j=1}^n\lambda_j|S_{ti}(\omega_j)|\Big/\sum_{k=1}^n\lambda_k\\&=\sum_{j=1}^n p_j|S_{ti}(\omega_j)|\frac{\lambda_j}{p_j\sum\limits_{k=1}^n\lambda_k}\\&\leqslant C\cdot E|S_{ti}|<\infty,\end{aligned}$$

其中 $C=\max\limits_{1\leqslant j\leqslant n}\frac{\lambda_j}{p_j}\left(\sum\limits_{k=1}^n\lambda_k\right)^{-1}<\infty$. 记 $S_{t0}=S_{00}(1+r)^t$，则有 $E^*(S_{ti}^*)<\infty$. 接下来只需要证明

$$E^*(S_{ti}^*-S_{si}^*|\mathcal{F}_s)=0,\quad P^*\text{-a. s.},$$

对 $s<t$ 以及任意的 $i=1,\cdots,d$ 成立即可. 实际上，由 P^* 的定义知，$S_{ti}^*-S_{si}^*$ 在$\mathcal{F}_s$ 下的条件期望为 0，如果对所有的 $A\in\mathcal{F}_s$，均有

$$0=\int_A 0\mathrm{d}P^*=\int_A(S_{ti}^*-S_{si}^*)\mathrm{d}P^*=E^*[\mathbf{1}_A(S_{ti}^*-S_{si}^*)].$$

153 则 P^* 就是鞅测度. ■

为了证明上述等式，令 $A\in\mathcal{F}_s$，定义策略 $\widetilde{\varphi}_u$ 如下

$$\widetilde{\varphi}_u=(\widetilde{\varphi}_{u0},\cdots,\widetilde{\varphi}_{ud})',\quad 0\leqslant u\leqslant T,$$

其中

$$\widetilde{\varphi}_{uj}(\omega)=\mathbf{1}_A(\omega)\mathbf{1}_{\{s+1\leqslant u\leqslant t\}}\mathbf{1}_{\{j=i\}}\quad \omega\in\Omega,\quad j=0,\cdots,d,$$

考虑对应的离散时间随机积分

$$\begin{aligned}Z_u&=\int_0^u\widetilde{\varphi}'_r\mathrm{d}S_r^*=\sum_{r=1}^u\widetilde{\varphi}'_r(S_r^*-S_{r-1}^*)\\&=\sum_{r=s+1}^{\min(u,t)}\mathbf{1}_A(S_{ri}^*-S_{r-1,i}^*),\quad u=0,\cdots,T.\end{aligned}$$

则 $Z_T=\mathbf{1}_A(S_{ti}^*-S_{si}^*)$. 显然 $(-Z_0,\ Z_T)\in U$，且 $\sum\limits_{\omega_j\in\Omega}\lambda_jZ_T(\omega_j)=0$. 因此，由 P^* 的定义知

$$E^*(\mathbf{1}_A(S_{ti}^*-S_{si}^*))=E^*(Z_T)=\sum_{j=1}^n\lambda_jZ_T(\omega_j)\Big/\sum_{k=1}^n\lambda_k=0.$$

从而 $\{S_t^*\}$ 是一个 P^* 鞅.

4.4 无套利市场的欧式未定权益

下面讨论在到期日 T 带有随机支付 X 的欧式未定权益 C，它依赖于原生资产. 首先介绍一下未定权益和衍生产品的区别.

定义 4.4.1 设 C 为一个非负随机变量，即 $C:\Omega\to[0,\ \infty]$ 是一个 $\mathcal{F}$-Borel 可测映射，则称 C 为一个**权益**或**未定权益**. 如果存在某个满秩可测函数 g，使得 $C=g(S_{Ti_1},\ \cdots,\ S_{Ti_j})$，则称未定权益 C 为原生资产 i_1，…，i_j，$1\leqslant j\leqslant d$ 的一个衍生产品. 称向量函数 $g=(g_1,\ \cdots,\ g_j)$ 为满秩函数，如果对任意的 $i=1$，…，j，不存在 Borel 可测函数 h 使得 $g_i=h(g_1,\ \cdots,\ g_{i-1},\ g_{i+1},\ \cdots,\ g_j)$.

定义 4.4.2 称欧式未定权益 C 为**可复制**或**可实现**的，如果存在交易策略 $\varphi=\{\varphi_t\}$，使得在到期日 T，其价值与 C 几乎必然相等，即 $C=\varphi'_TS_T$，P-a.s. 此时称 $\varphi=\{\varphi_t\}$ 为**复制策略**. 154

如 $\varphi=\{\varphi_t\}$ 为一个复制策略，$V_t=V_t(\varphi)=\varphi_t'S_t$，$1\leqslant t\leqslant T$ 为对应的价值过程，则有

$$C=V_T=V_0+\sum_{t=1}^T\varphi_t'(S_t-S_{t-1}),$$

等价地，有

$$C^*=V_T^*=V_0^*+\sum_{t=1}^T\varphi_t'(S_t^*-S_{t-1}^*).$$

反之亦然，即上述表达式表明 C 是可复制的，$\varphi=\{\varphi_t\}$ 是它的复制策略.

下列定理表明，一个可复制的未定权益的价值过程几乎必然是贴现未定权益的 Lévy 鞅. 该结论意味着计算贴现价值过程不必知道具体的复制策略.

定理 4.4.3 任一满足 $E(C)<\infty$ 的可复制未定权益 C 关于任一 P^* 是可积的. 且 C 的任一复制策略 φ 的贴现价值过程 $V_t^*=V_t^*(\varphi)$ 均为

$$V_t^* = E^*(C^* \mid \mathcal{F}_t), P^*\text{-a. s.}, \quad t = 1, \cdots, T.$$

特别地，V_t^* 在 P^* 下是一个 $\mathcal{F}_t$ 非负鞅.

证明 如果 V_t 的广义条件期望存在，则定理 4.2.8 成立，从而 $V_t\geqslant 0$，P-a. s. 是所证结论的一个充分条件. 由 $V_T=C\geqslant 0$，下面采用倒推法证明 $V_t\geqslant 0$，P-a. s. 设 $\varphi_t(c)=\varphi_t\mathbf{1}_{\{|\varphi_t|\leqslant c\}}$，其中 $c>0$，注意到

$$\varphi_t = \lim_{c\to\infty}\varphi_t(c), \quad X_{t-1} = \lim_{c\to\infty}X_{t-1}\mathbf{1}_{\{|\varphi_t|\leqslant c\}},$$

对 ω 几乎处处成立. 由

$$V_{t-1}^* = V_t^* - \varphi_t'(S_t^* - S_{t-1}^*) \geqslant -\varphi_t'(S_t^* - S_{t-1}^*)$$

及

$$\begin{aligned} E^*(V_{t-1}^*\mathbf{1}_{\{|\varphi_{t-1}|\leqslant c\}}) &= E^*(\mathbf{1}_{\{|\varphi_{t-1}|\leqslant c\}}\varphi_{t-1}'S_{t-1}^*) \\ &= E^*(\varphi_{t-1})(c)'S_{t-1}^*) \\ &\leqslant c\cdot\sum_{i=0}^{d}E^*|S_{ti}^*|<\infty \end{aligned}$$

得

$$\begin{aligned} V_{t-1}^*\mathbf{1}_{\{|\varphi_t|\leqslant c\}} &= E^*(V_{t-1}^*\mathbf{1}_{\{|\varphi_t|\leqslant c\}}\mid\mathcal{F}_{t-1}) \\ &\geqslant E^*(-\mathbf{1}_{\{|\varphi_t|\leqslant c\}}\varphi_t'(S_t^* - S_{t-1}^*)\mid\mathcal{F}_{t-1}) \\ &= -\varphi_t(c)'E^*(S_t^* - S_{t-1}^*\mid\mathcal{F}_{t-1}) \\ &= 0. \end{aligned}$$

取极限 $c\to\infty$ 得到 $V_{t-1}^*\geqslant 0$，P^*-a. s. 从而对所有的 t，$V_t\geqslant 0$，P-a. s. 因此，$E(V_t^*\mid\mathcal{F}_{t-1})$ 有定义.

在集合 $A_c=\{|\varphi_t|\leqslant c\}$ 上，有 $\varphi_t=\varphi_t(c)$，因而对任意 $c>0$，对几乎所有 $\omega\in A_c$ 有

$$\begin{aligned} E^*(V_t^*\mid\mathcal{F}_{t-1})(\omega) - V_{t-1}^*(\omega) &= E^*(V_t^* - V_{t-1}^*\mid\mathcal{F}_{t-1})(\omega) \\ &= E^*(\varphi_t(c)'(S_t^* - S_{t-1}^*)\mid\mathcal{F}_{t-1})(\omega) \\ &= \varphi_t(c)'E^*(S_t^* - S_{t-1}^*\mid\mathcal{F}_{t-1})(\omega) \\ &= 0. \end{aligned}$$

因此，由附录 A. 3. 1(x)知，$\mathbf{1}_{\Omega\setminus A_c}\cdot 0+\mathbf{1}_{A_c}\cdot V_{t-1}^*$ 是 $E(V_t^*\mid\mathcal{F}_{t-1})$ 的修正. 根据控制收敛定理，令 $c\to\infty$ 可得 $\{V_t^*\}$ 的一般鞅性质. 余下来，只需证明 $V_t^*=E(C^*\mid\mathcal{F}_t)$ 即可. 事实上，对 $t<T$，有

$$\begin{aligned} V_t^* &= E^*(V_{t+1}^*\mid\mathcal{F}_t) \\ &= E^*(\cdots E^*(V_T^*\mid\mathcal{F}_{T-1})\cdots\mid\mathcal{F}_t) \\ &= E^*(\cdots E^*(C^*\mid\mathcal{F}_{T-1})\cdots\mid\mathcal{F}_t) \end{aligned}$$

$$= E^*(C^* \mid \mathcal{F}_t),$$

证毕. ■

定义 4.4.4 设无套利市场有 d 个不同的风险资产和一个未定权益 C. 称 $\pi(C)\geqslant 0$ 为 C 的**无套利价格**，如果存在一个适应过程$\{S^*_{t,d+1}:t=0,\cdots,T\}$使得

(i) $S^*_{0,d+1}=\pi(C)$, P-a. s. ;

(ii) $S^*_{t,d+1}\geqslant 0$, P-a. s. , 对 $t=0,\cdots,T$;

(iii) $S^*_{T,d+1}=C^*$, P-a. s. ;

(iv) 由 $d+2$ 个价格过程$\{S_{t,0}\}$, …, $\{S_{t,d+1}\}$构成的延拓市场也是无套利的.

若可复制的未定权益 C 的交易价格 p 与对冲组合所需的初始资本 $V_0=\varphi_0' S_0=E^*(C^*)$ 不同，则套利者可从中赚取无风险利润. 例如，如果 $p>V_0$，投资者可以在 0 时刻卖出未定权益 C，而用 V_0 投资于对冲资产组合，则总能获得非负收益，其差价 $p-V_0>0$ 就是无风险利润. 在到期日 T，未定权益卖方只需通过该对冲组合的支付便可实现平仓.

156

定理 4.4.5 未定权益 C 的无套利价格集合为 $\Pi(C)=\{E^*(C^*):P^*\in\mathcal{P}, E^*(C^*)<\infty\}$.

证明 "$\subset$": 设 p 为一个无套利价格，则延拓后的市场是无套利的. 由定理 4.3.5 知 $\mathcal{P}\neq\varnothing$，因此存在 $P^*\in\mathcal{P}$，使得过程$\{S^*_{t,i}:t=0,\cdots,T\}(i=1,\cdots,d+1)$为 P^* 鞅. 下面证明 $p=E^*(C^*)$. 事实上，因为 $S_{T,d+1}=C$，由鞅的性质即有

$$p=S^*_{0,d+1}=E^*(S^*_{T,d+1}\mid\mathcal{F}_0)=E^*(C^*\mid\mathcal{F}_0)=E^*(C^*),$$

"$\supset$": 设 $p\in\{E^*(C^*):P^*\in P, E^*(C^*)<\infty\}$，则存在 $P^*\in\mathcal{P}$使得 $p=E^*(C^*)$. 固定 P^*，按如下方法构造 Lévy 鞅

$$S^*_{t,d+1}=E^*(C^*\mid\mathcal{F}_t),\quad t=0,\cdots,T.$$

由定义，$S^*_{t,d+1}$是$\mathcal{F}_t$ 适应的，且 $p=E^*(C^*)$. 由定理 4.4.3，$S^*_{t,d+1}$非负且几乎必然等于贴现价值过程. 因为 $d+1$ 维过程$\{S^*_{t,i}:t=0,\cdots,T\}(i=1,\cdots,d+1)$在 P^* 下是$\mathcal{F}_t$ 鞅，所以 P^* 是延拓后市场的等价鞅测度. 根据定理 4.3.5，P^* 是延拓后市场的等价鞅测度当且仅当延拓后的市场是无套利的. 所以，定义 4.4.4 的所有条件均已验证，证毕. ■

定义 4.4.6 称一个无套利的金融市场是**完备的**，如果它的所有未定权益 C 均是可复制的.

在一个完备的市场中，所有的未定权益都可以通过自融资交易策略实现对冲. 在第 2 章已经知道，单期模型中资产组合构成的线性空间与有界随机变量线性空间 $L_\infty(\Omega,\mathcal{F},P)$是一致的，它的维数不超过 $d+1$. 多期模型存在类似的结论.

定理 4.4.7 在一个无套利市场里，若所有的有界未定权益均为可实现的，则下列结论成立：

(i) 市场是完备的；

(ii) $L_\infty(\Omega,\mathcal{F},P)$的维数不超过$(d+1)^T$.

由定理 4.4.7 可得到：如果交易被严格限制在固定的离散时刻上，则任何一个无套利的完备市场均可由二叉树来表示. 在每个时刻 t，价值过程从最多 $d+1$ 个值的有限集合中取值. 换句话说，如果无套利市场不能用二叉树来表示，则存在不能被对冲的未定权益.

由定理 4.4.3 知，一个能被自融资交易策略对冲的未定权益的价值过程，满足下列关系

$$V_t^* = E^*(C^* \mid \mathcal{F}_t), \text{对所有 } P^* \in \mathcal{P} \text{成立}.$$

问题是，对不能对冲的未定权益，结论仍成立吗？上式右端对不同的 P^* 仍相同吗？ 157

定理 4.4.8 设 C 为一个欧式未定权益，则下列陈述等价.

（i）C 可由一个自融资策略复制.

（ii）对每一个等价鞅测度及 $t=0, \cdots, T$，条件期望 $E^*(C^* \mid \mathcal{F}_t)$，$P$-a.s. 有相同的值.

证明 （i）⇒（ii）它就是定理 4.4.3.

（ii）⇒（i）：假设 C 不能被复制，则 $C^* \notin \mathcal{V} = \text{span}\{\varphi' S_T^* : \varphi \in \mathbf{R}^{d+1}\}$，固定某个 $P^* \in \mathcal{P}$，令

$$(X, Y) = E^*(XY), \quad X, Y \in L_2(P^*).$$

作分解 $C^* = C_{\mathcal{V}}^* + C_{\mathcal{V}^\perp}^*$，其中 $C_{\mathcal{V}}^*(C_{\mathcal{V}^\perp}^*)$表示在$\mathcal{V}(\mathcal{V}^\perp)$上的正交投影. 定义

$$\widetilde{P}(\{\omega\}) = \left(1 + \frac{C_{\mathcal{V}^\perp}^*(\omega)}{\sup|C_{\mathcal{V}^\perp}^*|^2}\right) P^*(\{\omega\}).$$

因为 $\mathbf{1}_\Omega \perp C_{\mathcal{V}^\perp}^*$ $(\mathbf{1}_\Omega \in \mathcal{V})$，则有

$$\widetilde{P}(\Omega) = P^*(\Omega) + E^*\left(\frac{C_{\mathcal{V}^\perp}^*}{\sup|C_{\mathcal{V}^\perp}^*|^2} \mathbf{1}_\Omega\right) = P^*(\Omega) = 1,$$

进一步地，$P^* \sim \widetilde{P}$，$\widetilde{P}$是一个 EMM. 对任意 $Z \in \mathcal{V}$有

$$E_{\widetilde{P}}(Z) = E^*(Z) + \frac{1}{\sup|C_{\mathcal{V}^\perp}^*|^2} E^*(C_{\mathcal{V}^\perp}^* Z) = E^*(Z),$$

因此可得 $E_{\widetilde{P}}(C_{\mathcal{V}}^*) = E^*(C_{\mathcal{V}}^*)$. 另一方面

$$E_{\widetilde{P}}(C_{\mathcal{V}^\perp}^*) = E^*(C_{\mathcal{V}^\perp}^*) + \frac{1}{\sup|C_{\mathcal{V}^\perp}^*|^2} E^*(C_{\mathcal{V}^\perp}^{*2}) > E^*(C_{\mathcal{V}^\perp}^*).$$

两式相加即得

$$E_{\widetilde{P}}(C^* \mid \mathcal{F}_0) = E_{\widetilde{P}}(C^*) > E^*(C^*) = E^*(C^* \mid \mathcal{F}_0),$$

与已知矛盾. ■

根据定义，一个完备的金融市场是无套利的，从而至少可以找到一个 EMM. 还可证明：如市场是完备的，则只存在一个 EMM. 证明方法如下：若 $A \in \mathcal{F}$，则 $C = \mathbf{1}_A$ 是有界未定权益，且是可复制的，它的公平价格（即无套利价格）$E^*(C^*)$可由上面的定理唯一确定. 换句话说，映射

$$\Phi: \mathcal{P} \to \mathbf{R}, \quad P^* \mapsto E^*(\mathbf{1}_A) = P^*(A), \quad P^* \in \mathcal{P}$$

是一个常数，因此$\mathcal{P} = \{P^*\}$.

反过来，若$\mathcal{P} = \{P^*\}$，则集合$\{E^*(C^*) : P^* \in \mathcal{P}, E^*(C^*) < \infty\}$为无套利价格的单元素集，若 C 不可复制，就会出现多个无套利价格，与已知矛盾. 综上所述，有下列基本结论. 158

定理 4.4.9 无套利市场是完备的充分必要条件是存在唯一的等价鞅测度.

4.5 离散时间的鞅表示定理

本小节虽然篇幅较短，但内容却非常重要.

定理 4.5.1(鞅表示定理) 以下两个陈述等价.

(i) 市场中有且仅有一个等价鞅测度 P^*.

(ii) 设$\{M_t:t=0,\cdots,T\}$是一个 P^* 鞅，则$\{M_t:t=0,\cdots,T\}$可以表示成关于$\{S_t^*:t=0,\cdots,T\}$的随机积分，即存在某个可预报过程$\{\varphi_t:t=0,\cdots,T\}$，使得

$$M_t=M_0+\sum_{i=1}^{t}\varphi_i'(S_i^*-S_{i-1}^*)=M_0+\int_0^t\varphi_r'\mathrm{d}S_r^*.$$

证明 (i)⇒(ii)：设$\{M_t:t=0,\cdots,T\}$是一个 P^* 鞅，考虑分解式 $M_t=M_t^+-M_t^-$，M_t^+ 和 $M_t^-\geqslant 0$，其中 M_t^+、M_t^- 是贴现未定权益且 $EM_t^+<\infty$，$EM_t^-<\infty$. 因为$\mathcal{P}=\{P^*\}$，市场是完备的，则 M_t^+、M_t^- 可由自融资交易策略复制，因此，存在可测过程$\{\varphi_t^+\}$和$\{\varphi_t^-\}$使得$(\varphi_T^+)'S_T=M_T^+$，$(\varphi_T^-)'S_T=M_T^-$. 根据定理 4.4.3，贴现过程 $V_t^+=(\varphi_t^+)'S_t$，$V_t^-=(\varphi_t^-)'S_t$ 是 P^* 鞅，又因为

$$(V_t^+)^*=E^*((V_T^+)^*|\mathcal{F}_t)=E^*(M_T^+|\mathcal{F}_t),$$
$$(V_t^-)^*=E^*((V_T^-)^*|\mathcal{F}_t)=E^*(M_T^-|\mathcal{F}_t),$$

对所有 $t=0,\cdots,T$ 成立. 因此得到

$$\begin{aligned}M_t&=E^*(M_T^+-M_T^-|\mathcal{F}_t)\\&=(V_t^+)^*-(V_t^-)^*\\&=V_0^+-V_0^-+\sum_{j=1}^{t}(\varphi_i^+-\varphi_i^-)'(S_i^*-S_{j-1}^*),\end{aligned}$$

上述最后一个等式成立是因为交易策略$\{\varphi_t\}$对应的任意一个价值过程均可写成 $V_t=V_0+\sum_{i=1}^{t}\varphi_i'(S_i-S_{i-1})$.

(ii)⇒(i)：设 C 为一个未定权益，$E(C)<\infty$，考虑价值过程 $E(C|S_t^*)$，由定理 4.4.3 知它是一个 P^* 鞅. 根据假设，可以写成 $C_0+\int_0^t\varphi_r'\mathrm{d}S_r^*$，其中$\{\varphi_t\}$是某个可预报过程. 于是就得到了未定权益 C 的一个复制的交易策略. ■

159

值得一提的是，任何一个离散 P^* 鞅均可以表示成关于$\{S_t^*:t=0,\cdots,T\}$的随机积分. 该定理是数理金融中非常重要的一个结论. 事实上，上述证明过程中，每一项和每一步都有明确的经济含义，可以得出清晰的经济学解释.

4.6 Cox-Ross-Rubinstein 二叉树模型

著名的 Cox-Ross-Rubinstein 二叉树模型(简记为 CRR 模型)考虑这样的金融市场：假设市场上仅存在银行账户和一种风险资产，在每一期，风险资产的价格满足单期二叉树模型. 下面是其数学模型，令

$$\Omega = \{+,-\}^T = \{\omega = (\omega_1,\cdots,\omega_T)':\omega_t \in \{+,-\}, t = 1,\cdots,T\}.$$

于是，风险资产的价格过程$\{S_t: t=0, \cdots, T\}$有如下的形式

$$S_t(\omega) = \begin{cases} S_{t-1}(\omega)u, & \omega_t = +, \\ S_{t-1}(\omega)d, & \omega_t = -, \end{cases}$$

其中$\omega=(\omega_1, \cdots, \omega_T)'$，$t=1, \cdots, T$. S_0为初始股价，d和u分别为下降因子和上升因子，定义对应的收益

$$R_t = \frac{S_t - S_{t-1}}{S_{t-1}}, \quad t = 1,\cdots,T,$$

则收益取值于$\{r_+, r_-\}$，其中$r_+=u-1$，$r_-=d-1$. 注意，

$$\{R_t = r_+\} = \{\omega \in \Omega: \omega_t = +\} \text{ 和 } \{R_t = r_-\} = \{\omega \in \Omega: \omega_t = -\}.$$

考虑自然σ代数流：

$$\mathcal{F}_0 = \{\varnothing,\Omega\}, \quad \mathcal{F}_t = \sigma(S_0,\cdots,S_t), \quad t = 1,\cdots T,$$

设$\mathcal{F}=\mathcal{F}_T=\mathrm{Pot}(\Omega)$. 对了后面叙述方便，下面列举一些重要的公式

$$S_t = S_0 \prod_{i=1}^{t}(1+R_j), \quad S_t = S_{t-1}(1+R_t),$$

$$S_t^* = S_0^* \prod_{i=1}^{t}\frac{(1+R_i)}{1+r}, \quad S_t^* = S_{t-1}^*\frac{1+R_t}{1+r}.$$

其中，r是银行账户的固定利率.

下面的定理给出了CRR模型的无套利条件，并给出了等价鞅测度的显式公式.

定理 4.6.1 设$P(\{\omega\})>0$，$\forall\,\omega\in\Omega$.

(i) CRR模型是无套利的当且仅当

$$r_- < r < r+ \Leftrightarrow d < 1+r < u,$$

且在该条件下市场也是完备的.

160

(ii) 如果无套利条件$d<1+r<u$成立，则等价鞅测度P^*由下式给出

$$P^*(\{(\omega)\}) = (p^*)^k(1-p^*)^{T-k}, \quad k = 0,\cdots,T,$$

其中$\omega=(\omega_1, \cdots, \omega_T)'\in\Omega$, $k = k(\omega) = \sum_{t=1}^{T}\mathbf{1}_{\{\omega_t=+\}}$ 且

$$p^* = \frac{r-r_-}{r_+-r_-} = \frac{1+r-d}{u-d} \in (0,1).$$

进一步，有$\dfrac{dP^*}{dP}(\omega)=\dfrac{(p^*)^k(1-p^*)^{T-k}}{P(\{\omega\})}$，$\omega\in\Omega$. 且收益$R_1$，…，$R_T$在满足$P^*(R_1=r_+)=p^*$的$P^*$下是独立同分布的.

证明 假设模型是无套利的，则存在$P^*\in\mathcal{P}$使得S_t^*是一个P^*鞅. 由于$S_t^*=S_{t-1}^*\dfrac{R_{t+1}}{1+r}$，鞅条件$S_t^*=E^*(S_{t+1}^*\mid\mathcal{F}_t)$等价于$1+r=E^*(1+R_{t+1}\mid\mathcal{F}_t)$. 定义$P_t^*=P^*(1+R_{t+1}=1+r_+\mid\mathcal{F}_t)$，则上述方程可写为

$$1+r = (1+r_+)p_t^* + (1+r_-)(1-p_t^*)$$

$$= p_t^* (r_+ - r_-) + (1 + r_-).$$

因为 $r_- < r_+$，该方程有唯一解，解不依赖于时间，且对任意的 $t=1, \cdots, T$，

$$p^* = p_t^* = \frac{r - r_-}{r_+ - r_-} = \frac{1 + r - d}{u - d}.$$

注意到 $p^* \in [0, 1]$当且仅当 $r_- \leqslant r \leqslant r_+$，$p^* \in (0, 1)$当且仅当 $r_- < r < r_+$. 下面证明 P^* 的表达式，因为

$$\mathcal{F}_t = \sigma(S_0, \cdots, S_t) = \sigma(R_1, \cdots, R_t).$$

因此

$$P^*(R_{t+1} = r_+ \mid R_1, \cdots, R_t) = P^*(R_{t+1} = r_+ \mid \mathcal{F}_t) = p^*$$

表明，在已知 R_1，…，R_t 的条件下 R_{t+1} 的条件期望等于 $P^*(R_{t+1} = r_+)$，所以 R_{t+1} 独立于 R_1，…，R_t，从而 R_1，…，R_T 是在 P^* 下独立同分布的随机变量序列. 若定义 $f(+) = r_+$，$f(-) = r_-$，则有

$$\begin{aligned} P^*(\{\omega\}) &= P^*(R_1 = f(\omega_1), \cdots, R_T = f(\omega_T)) \\ &= (p^*)^k (1 - p^*)^{T-k} \end{aligned}$$

$\forall \omega = (\omega_1, \cdots, \omega_T)' \in \Omega$ 成立，所以 P^* 是唯一的，即$\mathcal{P} = \{P^*\}$. ■

需要指出的是，在等价鞅测度下收益序列是独立同分布的，但是该性质在实际概率测度下并不成立. 事实上，由统计分析可以得出金融市场上的收益序列是有相关性的. 直接计算可以得到 $E^*(R_1) = r_+ p^* + r_-(1 - p^*) = r$，这也表明，使用 P^* 可对未来支付进行风险中性定价.

161

利用 R_1，…，R_T 在 P^* 下是 i. i. d. 的结果，可直接得到下列资产价格分布的结果.

引理 4.6.2　在等价鞅测度下，

$$P^*(S_t = S_0 u^k d^{t-k}) = \binom{t}{k} (p^*)^k (1 - p^*)^{t-k},$$

对 $k=0$，…，t；$t=0$，…，T 成立.

现在可以容易地给出欧式未定权益的公平价格.

引理 4.6.3　设 $C = f(S_T)$为一个衍生品，其中 $f: \mathbf{R} \to [0, \infty)$是一个 Borel 函数. 则 C 的无套利价格为：

$$\pi(C) = \sum_{k=0}^{T} \frac{f(S_0 u^k d^{T-k})}{(1+r)^T} \binom{T}{k} (p^*)^k (1 - p^*)^{T-k}.$$

证明　只需注意到 $f(S_T)$取值 $f(S_0 u^k d^{T-k})$的概率为$\binom{T}{k}(p^*)^k(1-p^*)^{T-k}$即可. ■

CRR 模型的简单结构使得人们只需要知道模型的参数 u，d 和 p^*，就可以显性地计算出整个价值过程 V_t，V_t 显然是 S_t 的函数.

定理 4.6.4　令 $C = f(S_T)$为上述衍生品，则任何一个对冲策略的价值过程均满足

$$V_t = \sum_{k=0}^{T-t} \frac{f(S_t u^k d^{T-t-k})}{(1+r)^{T-t}} \binom{T-t}{k} (p^*)^k (1 - p^*)^{T-t-k}.$$

显然 V_t 是 S_t 的函数.

证明 记 $S_T = S_t \cdot \prod_{i=t+1}^{T}(1+R_i)$. 显然 S_t 是$\mathcal{F}_t$ 可测的，又因为 R_{t+1}，…，R_T 独立于 $\mathcal{F}_t$，所以

$$P^*\left(\prod_{i=t+1}^{T}(1+R_i) = u^k d^{T-t-k}\right) = \binom{T-t}{k}(p^*)^k(1-p^*)^{T-t-k}.$$

162

于是

$$\begin{aligned} V_t^* &= E^*(C^* \mid \mathcal{F}_t) \\ &= E^*\left(\frac{f(S_T)}{(1+r)^T}\middle|\mathcal{F}_t\right) \\ &= E^*_{(R_{t+1},\cdots,R_T)}\left(\frac{f\left(S_t\prod_{i=t+1}^{T}(1+R_i)\right)}{(1+r)^T}\middle|\mathcal{F}_t\right) \\ &= \sum_{k=0}^{T-t}\frac{f(S_t u^k d^{T-t-k})}{(1+r)^T}\binom{T-t}{k}(p^*)^k(1-p^*)^{T-t-k}. \end{aligned}$$

又因为 $V_t = (1+r)^t V_t^*$，所以结论成立. ■

人们甚至可以计算出路径相依的衍生品的价值过程，它 t 时刻的价值是依赖于到 t 时刻为止的股价 S_0，…，S_t 的函数 $v_t(S_0, \cdots, S_t)$.

引理 4.6.5 设存在 Borel 可测函数 $f:\mathbf{R}^{T+1}\to[0,\infty)$，使得 $C=f(S_0,\cdots S_T)$. 对 x_1，…，$x_t\geqslant 0$，定义

$$v_t(x_0,\cdots,x_t) = E^* f\left(x_0,\cdots,x_t,x_t(1+R_{t+1}),\cdots,x_t\prod_{i=t+1}^{T}(1+R_i)\right),$$

上式等价于

$$\int\cdots\int f\left(x_0,\cdots,x_t,x_t(1+r_{t+1}),\cdots,x_t\prod_{i=t+1}^{T}(1+r_i)\right)\mathrm{d}F^*(r_1)\cdots\mathrm{d}F^*(r_T),$$

其中 F^* 为 R_1 在 P^* 下的分布函数，则贴现价值过程为

$$V_t^* = E^*(C^* \mid \mathcal{F}_t) = \frac{v_t(S_0,\cdots,S_t)}{(1+r)^T}.$$

证明 记 $S_{t+j} = S_t \cdot \prod_{i=t+1}^{t+j}(1+R_i)$. 在 P^* 下，因子 $\prod_{i=t+1}^{T}(1+R_i)$ 独立于$\mathcal{F}_t=\sigma(S_0$，…，$S_t)$. 由 V_t 的定义有

$$\begin{aligned} (1+r)^T V_t^* &= E^*(C \mid \mathcal{F}_t) \\ &= E^*(f(S_0,\cdots,S_T \mid \mathcal{F}_t) \\ &= E^*(C(S_0,\cdots,S_T) \mid \mathcal{F}_t) \\ &= E^*\left(C\left(S_0,\cdots,S_t,S_t(1+R_{t+1}),\cdots,S_t\prod_{i=t+1}^{T}(1+R_i)\right)\right) \\ &= v_t(S_0,\cdots,S_t). \end{aligned}$$

因此，结论是显然的. ■

163

虽然上述公式是简单的，但是在实际应用中，二叉树模型的 T 却常常比较大，这是为了使得给定的区间能够充分离散化，或者是由于需要对一天的头寸进行多次调仓，比如，甚至每五分钟进行一次调仓而造成的. 在有大量的衍生品需要定价时，公式中的区间数目是影响计算成本的重要因素.

计算价值过程其实也是寻找复制的方法. 考虑序列 $S_0=V_0^*$，V_1^*，…，$V_T^*=C^*$，其中 V_t^* 由上面的公式计算. 要找一个可预报过程$\{\varphi_t\}$，使得它对应的价值过程复制上述值，也就是说要寻找到一个自融资交易策略 φ_t 使得对 $t=0$，…，$T-1$，有

$$V_t^* = V_{t-1}^* + \varphi_t(S_t - S_{t-1})$$

或，等价地

$$V_t^*(\omega) = V_{t-1}^*(\omega) + \varphi_t(\omega)(S_t(\omega) - S_{t-1}(\omega)), \quad \forall\, \omega \in \Omega.$$

在时刻 $t-1$，$V_{t-1}^*(\omega)$，$S_{t-1}^*(\omega)$都是已知的，$\varphi_t(\omega)$仅依赖于$(\omega_1, \cdots, \omega_T)$中的$(\omega_1, \cdots, \omega_{t-1})$，因而同样是固定的. $S_t(\omega)$的可能取值为 $S_{t,+}=S_t(+)$或 $S_{t,-}=S_t(-)$，$V_t(\omega)$的可能取值为 $V_{t,+}=V_t(+)$，$V_{t,-}=V_t(-)$，因此得到下列两个方程

$$V_{t,+}^* - V_{t-1}^* = \varphi_t(S_{t,+}^* - S_{t-1}^*)$$
$$V_{t,-}^* - V_{t-1}^* = \varphi_t(S_{t,-}^* - S_{t-1}^*),$$

当 $S_{t,+}^* \neq S_{t,-}^*$ 时，很容易得到该方程组的解.

定理 4.6.6 *设 C 是与路径有关的欧式衍生品，称下列自融资对冲组合为* **delta 对冲**

$$\varphi_t = \frac{V_{t,+}^* - V_{t,-}^*}{S_{t,+}^* - S_{t,-}^*} = \frac{V_{t,+} - V_{t,-}}{S_{t,+} - S_{t,-}}$$

(或简记为$\frac{\Delta V_t}{\Delta S_t}$). 它表示在 t 时刻，在标的资产上投资为 φ_t，而把余额 $V_{t-1}-\varphi_t S_{t-1}$ 以无风险利率存入银行或从银行借贷，是一种自融资对冲策略.

证明 在上述方程组中，由第一个方程减去第二个方程，消掉 V_{t-1}^* 和 S_{t-1}^*，得到

$$V_{t,+}^* - V_{t,-}^* = \varphi_t(S_{t,+}^* - S_{t,-}^*),$$

解得 φ_t 为 delta 对冲. 它表明，在时刻 $t-1$，如果平仓得到的收益为 $V_{t-1}=\varphi_{t-1}S_{t-1}$，而新的头寸成本为 $\varphi_t S_{t-1}$，因此，差价 $V_{t-1}-\varphi_t S_{t-1}$ 或者需从银行借贷，或者可以存入银行.

164

■

Delta 对冲是对冲衍生品的最关键的一个量. 因为在对冲组合中需要知道当前资产的组合情况，或者更准确地说，需要知道资产的变化情况，所以这里给出了这样一个重要的量.

4.7 Black-Scholes 公式

本节推导到期日为 T 的欧式看涨期权的无套利定价的 Black-Scholes 公式. 它是时间段为$[0, T]$，期数 $N\to\infty$ 的二叉树序列模型的极限情形. Black-Scholes 公式具有重要的意义. 首先，如果 N 太大而且有大量衍生品需要定价，由于计算量太大使用二叉树模型有困难；其次，该结果也表明增加 N 可以得到近以于极限的结果. 在 Black-Scholes 公式中，要求股价模型服从对数正态分布是一种自然且合适的选择.

因为，对 $\forall t \in \mathbf{N}$ 有 $S_t = S_0 \prod_{i=1}^{t}(1+R_i)$，其中 $R_i = \frac{S_i - S_{i-1}}{S_{i-1}}$，则有

$$\log(S_t) = \log(S_0) + \sum_{i=1}^{t} \log(1+R_i).$$

当 t 足够大，且收益是 i. i. d. 的时，根据中心极限定理，$\log(S_t)$是渐近正态的. 通过一些假定，可以得到股价的极限分布，设收益满足

$$\log(1+R_i) \overset{\text{i. i. d.}}{\sim} N(\mu, \sigma^2).$$

则有

$$\log(S_t) \sim N(\log(S_0) + \mu t, \sigma^2 t), \quad t \in \mathbf{N},$$

该分布在连续时间时仍然成立.

为了简化讨论，设 T 是一个正整数. 将$[0, T]$分割成为 $N = nT$ 个等距小区间，分点为

$$t_k = t_{nk} = \frac{k}{n}, \quad k = 0, \cdots, nT,$$

设在每个分点都有交易的可能性. 对 $n \in \mathbf{N}$，考虑 nT 期二叉树模型的第 n 期二叉树. 用 S_{nk}时间点 $k/n(k=0, \cdots, nT)$的股价. 对 $t \in [0, T]$，用与 t 最相近的交易时刻 t^* 的股价代替 t 时刻的股价，因为$\lfloor nt \rfloor / n \leqslant t \leqslant (\lfloor nt \rfloor + 1)/n$，所以 $t^* = \frac{\lfloor nt \rfloor}{n}$. 即定义 $S_n(t) = S_{n, \lfloor nt \rfloor}$，$t \in [0, T]$. 接下来，对固定 $t \in [0, T]$，讨论 $\log S_n(t)$，即

$$\log S_n(t) = \log(S_0) + \sum_{i=1}^{\lfloor nt \rfloor} \log(1+R_{ni}).$$

用 P_n^* 表示第 n 个节点对应的二叉树模型的等价鞅测度，下面将讨论在等价概率测度序列$\{P_n^*\}$下，$\log S_n(t)$是否依分布收敛，即 $S_n(t)$的极限分布是否为一个对数正态分布. 如果这种中心极限定理成立，则对 $\forall x \in \mathbf{R}$

$$F^*(x) = \lim_{n \to \infty} P_n^* \left(\frac{1}{\sqrt{n}} (\log S_n(t) - \log S_0) \leqslant x \right)$$

存在.

首先讨论如何确定第 n 期二叉树模型的参数 r_n，d_n，u_n. 设 r 为连续年利率，则初始时刻的每单位货币存入银行，到 T 时刻的值为 e^{rT}. 首先按下列方法选择第 n 期的固定利率 r_n，即要求二叉树模型得到与连续利率同样的累积收益，即

$$r_n = \mathrm{e}^{r/n} - 1,$$

由此得到$(1+r_n)^{nT} = (\mathrm{e}^{r/n})^{nT} = \mathrm{e}^{rT}$. 参数 d_n，u_n 必须满足无套利条件

$$0 < d_n < 1 + r_n < u_n, \quad n \in \mathbf{N}.$$

除此之外，可以限制 d_n，u_n 是一个重组的二叉树模型，即设

$$d_n u_n = 1, \quad n \in \mathbf{N},$$

因此可以选取 d_n，u_n 如下

$$u_n = \mathrm{e}^{\alpha_n}, \quad d_n = \mathrm{e}^{-\alpha_n}, \quad n \in \mathbf{N}.$$

165

其中$\{\alpha_n\}$为正数列. 注意到$\log S_n(t)$的被加数$\log(1+R_{ni})$在P_n^*下是 i. i. d. 的，其中

$$p_n^* = p_n^*((\log(1+R_{ni}) = \log(u_n)) = p_n^*(\log(1+R_{ni}) = \alpha_n).$$

因此，在P_n^*下，有

$$\log S_n(t) \stackrel{d}{=} \log(S_0) + \sum_{i=0}^{\lfloor nt \rfloor} Z_{ni},$$

其中Z_{n1}，…，Z_{nn}为取值于$\{-\alpha_n, \alpha_n\}$的随机变量列，且$P_n^*(Z_{n1}=\alpha_n)=p_n^*$. 由于标准中

166

心极限定理要求$\frac{1}{\sqrt{n}}\sum_{j=1}^{n}\xi_j$中的$\xi_1$，$\xi_2$，…是 i. i. d.，且是零期望和有限方差的. 因此，引入随机变量Y_{n1}，…，Y_{nn}，使其在P_n^*下是 i. i. d. 的，其中，

$$P_n^*(Y_{n1} = \sigma/\sqrt{n}) = 1 - P_n^*(Y_{n1} = -\sigma/\sqrt{n}) = p_n^*,$$

$\sigma>0$为某个常数，从而

$$\log S_n(t) \stackrel{d}{=} \log(S_0) + \frac{1}{\sqrt{n}}\sum_{i=1}^{\lfloor nt \rfloor} Y_{ni}.$$

综上所述，参数选择如下：

$$\begin{aligned} r_n &= e^{r/n} - 1 \\ d_n &= e^{-\sigma/\sqrt{n}}, \quad \text{使得 } r_{n-} = e^{-\sigma/\sqrt{n}} - 1 \\ u_n &= e^{\sigma/\sqrt{n}}, \quad \text{使得 } r_{n+} = e^{\sigma/\sqrt{n}} - 1. \end{aligned} \tag{4.3}$$

第n期二叉树模型对应的等价鞅测度由下式确定

$$p_n^* = \frac{r_n - r_{n-}}{r_{n+} - r_{n-}} = \frac{e^{r/n} - e^{-\sigma/\sqrt{n}}}{e^{\sigma/\sqrt{n}} - e^{-\sigma/\sqrt{n}}}. \tag{4.4}$$

p_n^*是否收敛于某个常数?

引理 4.7.1 下列结论成立.

(i) $\lim\limits_{n\to\infty} p_n^* = 1/2$.

(ii) $\lim\limits_{n\to\infty}\sqrt{n}(2p_n^* - 1) = \frac{r}{\sigma} - \frac{\sigma}{2}$.

证明 只需证明(ii). 注意到

$$\sqrt{n}(2p_n^* - 1) = \sqrt{n}\,\frac{2(e^{r/n} - 1) - (e^{\sigma/\sqrt{n}} + e^{-\sigma/\sqrt{n}} - 2)}{e^{\sigma/\sqrt{n}} - e^{-\sigma/\sqrt{n}}}.$$

由泰勒展开式$e^x = 1 + \sum_{k=1}^{\infty}\frac{x^k}{k!}$，

$$e^{r/n} - 1 = \frac{r}{n} + O\left(\frac{1}{n^2}\right),$$

$$e^{\sigma/\sqrt{n}} - 1 = \frac{\sigma}{\sqrt{n}} + \frac{1}{2}\frac{\sigma^2}{n} + \frac{1}{6}\frac{\sigma^3}{n^{3/2}} + O\left(\frac{1}{n^2}\right),$$

167

$$e^{-\sigma/\sqrt{n}} - 1 = -\frac{\sigma}{\sqrt{n}} + \frac{1}{2}\frac{\sigma^2}{n} - \frac{1}{6}\frac{\sigma^3}{n^{3/2}} + O\left(\frac{1}{n^2}\right),$$

得到

$$e^{\sigma/\sqrt{n}}+e^{-\sigma/\sqrt{n}}-2=\frac{\sigma^2}{n}+O\Big(\frac{1}{n^2}\Big),$$

$$e^{\sigma/\sqrt{n}}-e^{-\sigma/\sqrt{n}}=\frac{2\sigma}{\sqrt{n}}+\frac{1}{3}\frac{\sigma^3}{n^{3/2}}+O\Big(\frac{1}{n^2}\Big).$$

代入即有

$$\sqrt{n}(2p_n^*-1)=\frac{n}{\sqrt{n}}\frac{\frac{2r}{n}+O\Big(\frac{1}{n^2}\Big)-\Big(\frac{\sigma^2}{n}+O\Big(\frac{1}{n^2}\Big)\Big)}{\frac{2\sigma}{\sqrt{n}}+\frac{1}{3}\frac{\sigma^3}{n^{3/2}}+O\Big(\frac{1}{n^2}\Big)}$$

$$\xrightarrow{n\to\infty}\frac{2r-\sigma^2}{2\sigma}=\frac{r}{\sigma}-\frac{\sigma}{2}.$$

注意，在一个重组的二叉树模型中，价格过程的路径由价格上升的次数确定，而价格上升的次数服从二项分布．下列形式的中心极限定理将被用到． ■

定理 4.7.2 令 Y_n，$n\in\mathbf{N}$ 是服从二项分布的随机变量序列，参数为 n，$p_n\in(0,1)$，且$\lim\limits_{n\to\infty}p_n=p\in(0,1)$，则有

$$\frac{Y_n-np_n}{\sqrt{np_n(1-p_n)}}\xrightarrow[n\to\infty]{d}N(0,1).$$

证明 记$Y_n=\sum_{i=1}^{n}X_{ni}$，其中 X_{n1}，…，$X_{nn}\overset{\text{i.i.d.}}{\sim}\text{Bin}(p_n)$．显然随机变量 $X_{ni}-p_n(i=1,\cdots,n)$的值不超过 2，且具有零均值．进一步，$s_n^2=\text{Var}(Y_n)=np_n(1-p_n)\to\infty$，$n\to\infty$．因此满足$\delta=1$的李雅普诺夫不等式条件，则当 $n\to\infty$时，

$$\frac{1}{s_n^3}\sum_{i=1}^{n}E|X_{ni-}p_n|^3\leqslant\frac{2}{s_n^3}\sum_{i=1}^{n}E(X_{ni-}p_n)^2$$

$$\leqslant\frac{2}{S_n}\to 0.$$ ■ 168

定理 4.7.3（二叉树模型依分布收敛） 在(4.4)式给出的等价鞅测度$\{P_n^*\}$下，如果模型参数由(4.3)式给出，则当 $n\to\infty$时，对任意一个 t，有

$$S_n(t)\xrightarrow[n\to\infty]{d}S_t=S_0\exp\Big(\sigma B_t+\Big(r-\frac{\sigma^2}{2}\Big)t\Big),$$

其中，$B_t\sim N(0,t)$．等价于在极限概率测度 P^* 下，$S_n(t)$服从渐近对数正态分布，即

$$\log S_n(t)n\xrightarrow[n\to\infty]{d}\log(S_0)+\Big(r-\frac{\sigma^2}{2}\Big)t+\sigma B_t. \tag{4.5}$$

证明 显然有

$$S_n(t)=S_0u_n^{K_{\lfloor nt\rfloor}}d_n^{\lfloor nt\rfloor-K_{\lfloor nt\rfloor}}$$

其中 $K_{\lfloor nt\rfloor}\sim\text{Bin}(\lfloor nt\rfloor,p_n^*)$是到时刻$\lfloor nt\rfloor$为止股价上升的次数．因为 $f(z)=e^z$，$z\in\mathbf{R}$ 是连续函数，只需证明(4.5)式即可．由定义有 $\log u_n=-\log d_n=\frac{\sigma}{\sqrt{n}}$，所以

$$\log S_n(t)=\log(S_0)+\frac{2\sigma}{\sqrt{n}}K_{\lfloor nt\rfloor}-\frac{\sigma\lfloor nt\rfloor}{\sqrt{n}}.$$

在上式中标准化 $K_{\lfloor nt\rfloor}$得到

$$\log S_n(t)=\log(S_0)+\frac{2\sigma\sqrt{\lfloor nt\rfloor p_n^*(1-p_n^*)}}{\sqrt{n}}\cdot\frac{K_{\lfloor nt\rfloor}-\lfloor nt\rfloor p_n^*}{\sqrt{\lfloor nt\rfloor p_n^*(1-p_n^*)}}$$
$$+\frac{2\sigma\lfloor nt\rfloor p_n^*}{\sqrt{n}}-\frac{\sigma\lfloor nt\rfloor}{\sqrt{n}}.$$

其中第三和第四项可以简化为 $\sigma\frac{\lfloor nt\rfloor}{n}\sqrt{n}(2p_n^*-1)$，由引理 4.7.1，它收敛于 $t(r-\sigma^2/2)$；又因为 $p_n^*\to\frac{1}{2}$，$n\to\infty$，由定理 4.7.2 知，当 $n\to\infty$时

169

$$\frac{K_{\lfloor nt\rfloor}-\lfloor nt\rfloor p_n^*}{\sqrt{\lfloor nt\rfloor P_n^*(1-p_n^*)}}\xrightarrow[n\to\infty]{d}N(0,1),$$

从而得到

$$\lim_{n\to\infty}\frac{2\sigma\sqrt{\lfloor nt\rfloor p_n^*(1-p_n^*)}}{\sqrt{n}}=\sigma\sqrt{t}.$$

记 $U\sim N(0,1)$，当 $n\to\infty$，由 Slutzky 引理得

$$\log S_n(t)\xrightarrow{d}\log(S_0)+\sigma\sqrt{t}\cdot U+\left(r-\frac{\sigma^2}{2}\right)t,$$

令 $B_t=\sqrt{t}U\sim N(0,t)$即为所需结论. ■

上面的极限分布称为几何布朗运动. 定理 4.7.3 表明，在每个固定时间 t，股价的极限分布是一个几何布朗运动. 事实上，还可以在随机(右连左极)函数弱收敛的意义下得到这个收敛结果，见附录 B.2.

以上结论表明，在连续时间中，股价模型可以表示为

$$S_t=S_0\exp\left(\sigma B_t+\left(r-\frac{\sigma^2}{2}\right)t\right),\quad t\in[0,T].$$

下面就可以很自然地推导出著名的 Black-Scholes 公式. 显然，贴现股价

$$S_t^*=S_t\mathrm{e}^{-rt}=S_0\exp\left(\sigma B_t-\frac{\sigma^2}{2}t\right),t\in[0,T]$$

服从对数正态分布. 注意到看跌期权的支付函数 $f(x)=\max(K-x,0)$是一个以 K 为界的连续函数，由定理 4.7.3 知

$$E^*(\max(K-S_n(T),0)\mathrm{e}^{-rT})\to_{n\to\infty}E^*(\max(K-S_T,0)\mathrm{e}^{-rT}).$$

根据看跌-看涨平价关系，若看涨期权的支付函数为 $C(S_T)=\max(S_T-K,0)$，则有

$$E^*(\max(K-S_n(T),0)\mathrm{e}^{-rT})=E^*(C(S_T)\mathrm{e}^{-rT})-S_0+K\mathrm{e}^{-rT},$$

从而也得到欧式看涨期权的无套利价格的收敛结果

$$E^*(C(S_T)\mathrm{e}^{-rT})\to_{n\to\infty}E^*(\max(S_T-K,0)\mathrm{e}^{-rT}),$$

上式右边使用 1.5.5 小节的计算结果，从而得到

$$E^*(C(S_T)e^{-rT}) = S_0\Phi(d_1) - Ke^{-rT}\Phi(d_2),$$

其中 d_1，d_2 的定义已在 1.5.5 节给出. 170

4.8 美式期权和美式未定权益

现在讨论美式未定权益的定价问题，美式未定权益在到期日之前的任意时刻均可执行交易. 利用第 3 章讨论的最优停止理论，本节将讨论对冲策略的构造及美式期权无套利价格的计算.

在实际中，美式未定权益也经常在二叉树模型下进行定价，下面给出详细定价方法及具体算法.

4.8.1 无套利定价和期权执行策略

美式未定权益的持有者可以在到期日 T 之前的任意时刻 $t=0, \cdots, T$ 执行交易. 设在时刻 t 投资者收到支付 F_t，考虑基于股价 S_t 的美式看涨期权，其内在价值(或不执行时的价值)为 $S_t - K$，对应的支付为

$$F_t = \max\{S_t - K, 0\},$$

如果 $S_t \leqslant K$，期权值为 0. 以下所得的大部分结果都适用于任意的美式衍生品，甚至任意的权益. 记

$$F_t = f(S_t), \quad t = 0, \cdots, T,$$

其中 f 是可测函数，$\{F_t\}$是$\mathcal{F}_t$ 适应过程. 本节需要解决如下问题：

(i) 美式未定权益的公平价格是多少?

(ii) 何时行权可以最大化期望价值?

设金融市场是无套利且完备的，$\{\varphi_t\}$是美式未定权益的价值过程$\{V_t\}$的对冲策略，使得

$$V_t = V_t(\varphi), \quad t = 0, \cdots, T.$$

则所求的无套利价格 x 就是构建对冲组合所需的初始成本. 显然有

$$V_0 = x$$

$$V_t \geqslant F_t, \quad t = 1, \cdots, T.$$

对于一个完美的对冲组合，上述不等式应该为等式. 设 P^* 为等价鞅测度，贴现价值过程

$$M_t = V_t^* = e^{-rt}V_t, \quad t = 0, \cdots, T$$ 171

在 P^* 下是一个$\mathcal{F}_t$ 鞅. 由定理 3.2.18 知，对任意的停时 τ，有

$$E^*(V_\tau^*) = M_0 = V_0. \tag{4.6}$$

因为对冲$\{\varphi_t\}$保证了 $\forall t$，$V_t \geqslant F_t$，所以对 $\forall t$，有结论 $V_t^* = e^{-rt}V_t \geqslant e^{-rt}F_t = F_t^*$.

因此，$V_\tau^* \geqslant F_\tau^*$ 对任意停时 $\tau: \Omega \to \mathbf{N}_0$ 成立，结合(4.6)式有

$$V_0 = E^*(e^{-r\tau}V_\tau) \geqslant E^*(e^{-r\tau}F_\tau).$$

特别地，无套利价格满足

$$x \geqslant E^*(e^{-r\tau}F_\tau)$$

对任意停时 $\tau\in\mathcal{T}[0, T]$均成立，其中

$$\mathcal{T}[0,T]=\{\tau:\tau\text{ 是停时且 }\tau(\Omega)\subset\{0,\cdots,T\}\}$$

为最迟在 T 时刻必须停止的所有停时的集合. 换句话说，如果对冲是完美的，则

$$x\geqslant\sup_{\tau\in\mathcal{T}_{0,T}}E^*(F^*_{\tau^*})$$

为等式. 问题变为需要确定最优停时问题的最优值，或寻找 $\tau^*\in\mathcal{T}_{0,T}$，使 $E^*(F^*_{\tau^*})$最大. 这个问题在 3.2.2.3 节中已经解决了. 下一步需要确定$\{F^*_t:t=0,\cdots,T\}$的 snell 包络，它由如下表达式递归地定义：

$$Z_T=F^*_T,$$

$$Z_t=\max\{F^*_t,E^*(F^*_{t+1}\mid\mathcal{F}_t)\},\quad t=0,\cdots,T-1.$$

根据定理 3.2.23，最优停时就是贴现支付 F^*_t 不小于 $E^*(F^*_{t+1}\mid\mathcal{F}_t)$的首达时间，其中，对 P-几乎所有状态 $\omega\in\Omega$，$F^*_t=Z_t$. 换句话说，

$$\tau^*=\min\{t\geqslant 0:Z_t=F^*_t\}$$

满足

172

$$Z_0E^*(Z_{\tau^*})=\sup_{\tau\in\mathcal{T}_{0,T}}E^*(Z_\tau).$$

因此，得到的最优价值进而无套利价格为 $x=Z_0$. 由于对冲组合必须保证 $V^*_t\geqslant F^*_t$，而完美对冲组合使等式成立，因此 V^*_t 已经是一个 snell 包络. 事实上，该结论成立可由下列命题得出，命题的证明给出了实际中构造对冲组合的方法.

命题 4.8.1 一个美式未定权益的贴现支付 F^*_t($t=0,\cdots,T$)的 snell 包络由一个完美对冲的贴现价值过程 V^*_t($t=0,\cdots,T$)给出.

证明 由定义有 $F_T=V_T$. 在时刻 $T-1$，如果行权则得到支付 F_{T-1}，或者继续持有直至 T 时刻. 在定理 4.4.3 中取 $C^*=F^*_T$，则未定权益在 $T-1$ 时刻的贴现价值 V^*_{T-1}可由条件期望 $E^*(F^*_T\mid\mathcal{F}_{T-1})$表示. 因此，$T-1$ 时刻不行权的价值为

$$N_{t-1}=\mathrm{e}^{r(T-1)}E^*(F^*_T\mid\mathcal{F}_{T-1}).$$

注意到，N_{T-1}也是到期日为 T，支付为 F_T 的单期模型的看涨期权的无套利价格，人们可以用该组合来复制期权以实现对冲. 进一步，当知道了 N_{T-1}，若 $F_{T-1}\geqslant N_{T-1}$，则可以提前行权；若 $F_{T-1}\leqslant N_{T-1}$，则继续持有头寸至到期日. 因此，对冲必须保证

$$V_{T-1}=\max\{F_{T-1},\mathrm{e}^{r(T-1)}E^*(F^*_T\mid\mathcal{F}_{T-1})\},$$

其贴现形式为：

$$V^*_{T-1}=\max\{F^*_{T-1},E^*(F^*_T\mid\mathcal{F}_{T-1})\}.$$

利用同样的方法，逐步后推得到，对 $t=0,\cdots,T-1$，

$$V_t=\max\{F_t,\mathrm{e}^{-r}E^*(F_{t+1}\mid\mathcal{F}_t)\}$$

贴现形式为

$$V^*_t=\max\{F^*_t,E^*(F^*_{t+1}\mid\mathcal{F}_t)\},$$

其中用到了 $F^*_{t+1}=\mathrm{e}^{-r(t+1)}F_{t+1}$，从而

$$\mathrm{e}^{rt}E^*(F^*_{t+1}\mid\mathcal{F}_t)=\mathrm{e}^{-r}E^*(F_{t+1}\mid\mathcal{F}_t).$$

所以，$\{V_t^*\}$是$\{F_t^*\}$的 snell 包络．上述过程表明，通过倒向递推的方式可以构造美式未定权益的对冲组合． ■ 173

4.8.2 美式期权的二叉树定价

前面已经讨论过，在 N 期二叉树模型中，股价在一期或上升或下降，即

$$S_t(\omega)=\begin{cases}S_{t-1}u, & \omega_t=+,\\ S_{t-1}d, & \omega_t=-,\end{cases}$$

其中 $\omega=(\omega_1, \cdots, \omega_N)\in\Omega=\{+, -\}^N$，$u$ 是上升因子，d 是下降因子．如果时间步长为 Δ，则等价鞅测度为

$$p^*=\frac{e^{r\Delta}-e^{-\sigma\Delta}}{e^{\sigma\Delta}-e^{-\sigma\Delta}},$$

其中 $\sigma>0$ 为常数．在 4.7 节中，令 $\Delta=1/\sqrt{n}$，其中 n 为时间区间（通常指一年）的分割数目．由于模型中有 n 个节点，记为$(n, 1)$，…，(n, n)，它们由股价的变化次数唯一地确定．引入记号

S_{ni}：股票在(n,i)时的价格

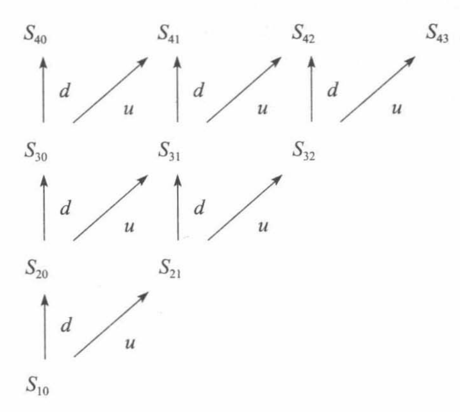
F_{ni}, F_{ni}^*：期权在(n,i)行权时的（贴现）支付

N_{ni}：(n,i)不行权时的价值

V_{ni}, V_{ni}^*：期权在(n,i)时的（贴现）价值

E_{ni}：期权在(n,i)时行权，取值为 1；否则为 0

现在按照自然顺序来排列二叉树的节点，使得节点(n, i)对应 i 个上升，下面的图片给出了这种形式：



根据模型，对 $i=0$，…，$n-1$；$n=1$，…，N，有 $S_{ni}=S_0d^iu^{n-i}$和 $F_{ni}=f(S_0d^iu^{n-i})$成立，而后者等于 $F_{ni}=\max\{S_0d^iu^{n-i}-K, 0\}$．如果未定权益为美式看涨期权，则在节点$(n, i)$处不行权时的价值为 174

$$N_{ni}=e^{-r\Delta}(p^*F_{n+1,j+1}+(1-p^*)F_{n+1,j}),$$

从而

$$V_{ni} = \max\{F_{ni}, e^{-rt}(p^* F_{n+1,j+1} + (1-p^*)F_{n+1,j})\}.$$

而如果 $F_{ni} \geqslant e^{-rt}(p^* F_{n+1,j+1} + (1-p^*)F_{n+1,j})$，则令 $E_{ni}=1$，此时投资者会在节点(n, i)处行权．二叉树计算一般从叶节点开始，一级级往回算，直到根节点，最后期权的价格即为 V_0．

4.9 评注与延伸阅读

关于离散时间的套利理论，权威性的著作是 Follmer 和 Schied(2004)，它包含了本章的全部内容，特别地，它给出了本章定理 4.4.7 的详细证明，而本章定理 4.3.3 是它中命题 5.11 的简化形式．全面介绍套利理论的著作还有 Shiryaev(1999)、Bingham 和 Kiesel (2004)．要对离散时间的套利理论作初等性的了解，不妨参考 Pliska(1997)和 Shreve (2004)．多期二叉树模型的工作起源于 Cox 等(1976)，在精确的闭型公式尚未发现的情况下，他给出了实物期权定价的主要框架．要了解更多关于建立在 Esscher 变换基础上多期模型的鞅测度理论，可以参考 Shiryaev(1999)．

参考文献

Bingham N.H. and Kiesel R. (2004) *Risk-Neutral Valuation: Pricing and Hedging of Financial Derivatives*. Springer Finance 2nd edn. Springer-Verlag London Ltd., London.

Cox J.C., Ross S.A. and Rubinstein M. (1976) *Option pricing: A simplified approach*. Journal of Financial Economics.

Föllmer H. and Schied A. (2004) *Stochastic Finance: An Introduction in Discrete Time*. vol. 27 of *de Gruyter Studies in Mathematics* extended edn. Walter de Gruyter & Co., Berlin.

Pliska S. (1997) *Introduction to Mathematical Finance*. Blackwell Publishing, Oxford.

Shiryaev A.N. (1999) *Essentials of Stochastic Finance: Facts, Models, Theory*. vol. 3 of *Advanced Series on Statistical Science & Applied Probability*. World Scientific Publishing Co. Inc., River Edge, NJ. Translated from the Russian manuscript by N. Kruzhilin.

Shreve S.E. (2004) *Stochastic Calculus for Finance. I: The Binomial Asset Pricing Model*. Springer Finance. Springer-Verlag, New York.

175

第5章　布朗运动和相关的连续时间过程

本章主要介绍连续时间随机过程的基本知识，特别是布朗运动及其性质和一些相关的过程. 对布朗运动的研究是现代科学的重要部分，布朗运动是由植物学家 Robert Brown 于 1827 年凭经验发现的. 布朗在研究微粒子在流体中的扩散时，发现它们以不规则的方式运动. 1905 年 Albert Einstein 发表了一个物理定理，得到了这种运动的数学方程及其解的计算方法，解释了这种被许多更小的流体分子冲撞形成的不规则运动. 1906 年，Marian Smoluchowski 发表了一个类似的结果. 其实早在 1900 年，Louis Bachelier 就使用了布朗运动的一维形式 $t \mapsto B_t (t \geqslant 0)$ 去建立股票价格模型. 这种经济推理是类似的，买卖股票的竞价流导致股价的上下波动. Norbert Wiener 建立了布朗运动的严格数学理论，他首次证明了布朗运动的存在性，因此布朗运动又称为维纳过程. 维纳还引入了维纳测度，用以刻画作为时间 t 的函数的布朗运动的分布.

布朗运动有几个特殊的性质，乍一看，是相当令人费解的：路径是几乎必然连续的，但处处不可微；在任何区间 $[a, b](a<b)$ 上，曲线 $t \mapsto B_t (t \in [a, b])$ 的长度都是无限的. 本章将仅限于讨论连续时间随机过程的一些非常有用的知识点，包括连续时间鞅(半鞅、上鞅)的定义的扩展、布朗运动的严格定义及它的一些重要的基本性质与计算法则. 除此之外，本章还将介绍一些相关的过程，比如布朗桥、几何布朗运动及 Lévy 过程.

5.1　预备

设 $I \subset [0, \infty]$ 是一个区间，例如，$I=[0, T]$.

定义 5.1.1　设 $d \in \mathbf{N}$,

$$X_t : (\Omega, \mathcal{F}, P) \to (\mathbf{R}^d, \mathcal{B}^d), \quad t \in I,$$

则称随机变量($d=1$)或随机向量($d>1$)族$\{X_t, t \in I\}$为**连续时间的随机过程**.

在本书中，记号 $X_t(\omega)$ 与 $X(t, \omega)$ 表示相同的含义，而且在指标集 I 不至于引起混淆的情况下，将$\{X_t, t \in I\}$简记为$\{X_t\}$.

将随机过程看作时间变量 t 的函数，不论是在应用中还是在理论上都是至关重要的.

定义 5.1.2

（ⅰ）在随机过程$\{X_t, t \in I\}$中，对固定的 $\omega \in \Omega$，函数 $t \mapsto X(t, \omega)$ 称为 X 的一条(简单)**路径**或轨迹，或一个**实现**.

（ⅱ）如果对任意给定的 $\omega \in \Omega$，函数 $t \mapsto X(t, \omega)$ 是左(右)连续的或连续的，则称随机过程$\{X_t\}$是左(右)连续的或连续的. 如上述条件仅在 $\omega \in \Omega \setminus N$ 上成立，其中 N 是一个零测集，则称$\{X_t\}$是 a. s.(几乎必然)左(右)连续，或 a. s.(几乎必然)连续的.

（ⅲ）一个 a. s. 右连续且左极限存在的过程称为一个 **càdlàg**(右连左极)过程.

càdlàg 术语是法语 continu à droite，limite à gauche 的首字母简写，意思是右连续且左极限存在. 下列概念是离散时间向连续时间的直接推广.

定义 5.1.3

(ⅰ) 如果对任意 $s, t\in I$，$s<t$ 均有

$$\mathcal{F}_s\subset\mathcal{F}_t\subset\mathcal{F},$$

则称子 σ 代数族$\{\mathcal{F}_t, t\in I\}$为滤子.

(ⅱ) 如果对每个 $t\in I$，X_t 都是$\mathcal{F}_t$ 可测的，则称随机过程$\{X_t, t\in I\}$是$\mathcal{F}_t$ **适应的**.

(ⅲ) 如果滤子$\mathcal{F}_t=\sigma(X_s:s\leqslant t)=\sigma(\bigcup_{s\leqslant t}X_s^{-1}(\mathcal{B}^d))$，则称该滤子为**自然滤子**.

(ⅳ) 引入滤子$\{\mathcal{F}_t\}$后的概率空间$(\Omega, \mathcal{F}, P)$称为**过滤概率空间**，记为$(\Omega, \mathcal{F}\{\mathcal{F}_t\}, P)$.

定义 5.1.4 设$\{M_t, t\in I\}$是一个$\mathcal{F}_t$ 适应的随机过程，且

178

(ⅰ)
$$E|M_t|<\infty,\quad \forall t\in I.$$

那么，

(ⅱ) 如果对任意 $s, t\in I$，且 $s\leqslant t$，均有 $M_s=E(M_t|\mathcal{F}_s)(P\text{-a. s.})$，则称$\{M_t, t\in I\}$为一个**鞅**.

(ⅲ) 如果对任意 $s, t\in I$，且 $s\leqslant t$，均有 $M_s\leqslant E(M_t|\mathcal{F}_s)(P\text{-a. s.})$，则称$\{M_t, t\in I\}$为一个**下鞅**.

(ⅳ) 如果对任意 $s, t\in I$，且 $s\leqslant t$，均有 $M_s\geqslant E(M_t|\mathcal{F}_s)(P\text{-a. s.})$，则称$\{M_t, t\in I\}$为一个**上鞅**.

下面给出一些例子.

▶**例 5.1.5**

(ⅰ) 设 A 是定义在$(\Omega, \mathcal{F})$上的随机变量，令

$$X_t(\omega)=X(t,\omega)=A(\omega)\cdot\sin(t),\quad t\in[0,2\pi],\quad \omega\in\Omega.$$

那么，$\{X_t:t\in[0, 2\pi]\}$是一个路径连续过程，即是一个连续过程.

(ⅱ) 设 $X_1, X_2, \cdots, X_n$ 是一个$(\Omega, \mathcal{F}, P)$上的随机样本. 考虑经验分布函数

$$F_n(t)=\frac{1}{n}\sum_{i=1}^{n}\mathbf{1}(X_i\leqslant t),\quad t\in\mathbf{R}.$$

则$\{F_n(t):t\in\mathbf{R}\}$是一个右连左极过程，因为当 ω 固定时，所有路径 $t\mapsto F_n(t, \omega)$都是右连续的阶梯函数.

(ⅲ) 对过程$\{X_t:t\in[0, T]\}$，考虑自然滤子$\mathcal{F}_t=\sigma(X_s:s\leqslant t)$，$t\in[0, T]$，那么，过程

$$Y_t:=X_t^2-t,\quad t\in I,$$

是$\mathcal{F}_t$ 适应的.

(ⅳ) 设$\{X_t:t\in I\}(I=[0, T])$是一个连续过程，令$\mathcal{F}_t=\sigma(X_s:s\leqslant t)$. 因为对每一个固定的 ω，$s\mapsto X_s(\omega)$都是连续的，从而对每一个 $\omega\in\Omega$，可以定义过程

$$Z_t(\omega)=\int_0^t X_s(\omega)\mathrm{d}s,\quad t\in I,$$

式中的积分为 Riemann 积分. 由构造可看出$\{Z_t:t\in I\}$是一个$\mathcal{F}_t$ 适应过程. 在下一章中，

我们还将更详细地讨论这个随机积分. ◀

在后面的章节中，我们不再强调可测性问题，但请注意，前面的定义中不需要函数 $t\mapsto X_t(\omega)(t\in I)$ 和 $X(\cdot,\cdot):I\times\Omega\to\mathbf{R}$ 是可测的. 如果后者是可测的，则称随机过程 X 是**可测的**，其中，$I\times\Omega$ 上的 σ 代数为 $\mathcal{B}(I)\times\mathcal{F}$，$\mathbf{R}^d$ 中的 σ 代数就是其 Borel-σ 代数 $\mathcal{B}^d$. 还请记住，在经济应用领域中适应过程很重要，这就自然要求在每一个给定的时刻 t，对 $\forall s\in[0,t]$，$\omega\in\Omega$，映射 $X(s,\omega)$ 关于到 t 时刻为止的所有事件集可测. 这就引出了下列定义：如果对任一固定的 t，映射 [179]

$$X:[0,t]\times\Omega\to\mathbf{R}^d,\quad (s,\omega)\mapsto X(s,\omega),\quad s\in[0,t],\quad \omega\in\Omega$$

是 $\mathcal{B}([0,t])\times\mathcal{F}_t$ 可测的，其中 $\mathcal{B}([0,t])\times\mathcal{F}_t$ 是 $[0,t]\times\Omega$ 上的 σ 代数，则称过程 X 是**循序可测的**. 因为 $\mathcal{B}([0,t])\times\mathcal{F}_t\subset\mathcal{B}(I)\times\mathcal{F}$，故循序可测过程必为可测过程.

在实际应用中，右连左极过程是非常重要的，下面的结果确保了这类过程在可测性上不存在问题.

引理 5.1.6 *任一左连续或右连续的适应过程都是循序可测的.*

证明 对右连续适应过程 $X=\{X_t\}$，定义

$$X_{nt}(s,\omega)=\sum_{i=1}^{n}X(it/n,\omega)\mathbf{1}_{((i-1)t/n,it/n]}(s).$$

固定 t，注意到对任一 Borel 集 A，

$$\{0\}\times\{\omega:X(0,\omega)\in A\}$$

与

$$\bigcup_{i=1}^{n}((i-1)t/n,it/n]\times\{\omega:X(it/n,\omega)\in A\}$$

构成

$$E_t=\{(s,\omega):0\leqslant s\leqslant t,\omega\in\Omega,X_{nt}(s,\omega)\in A\}$$

的一个分割. 由于 X 是适应的，从而 $E_t\in\mathcal{B}([0,t])\times\mathcal{F}_t$，这表明 X_{nt} 是循序可测的. 根据右连续性可知，当 $n\to\infty$ 时，对 $\forall s\in[0,T]$，$\omega\in\Omega$，当 $n\to\infty$ 时，在点态意义下有

$$X_{nt}(s,\omega)\to X(s,\omega),$$

因为可测函数的极限仍是可测的，这就证明了 X 是循序可测的.

如 (X,Y) 是二维随机向量，且 $EX^2<\infty$，$EY^2<\infty$，那么下面的一些量存在：

$$\mu_X=E(X),\quad \mu_Y=E(Y),\quad \sigma_X^2=\mathrm{Var}(X),\quad \sigma_Y^2=\mathrm{Var}(Y)$$

$$\rho_{XY}=\mathrm{Cov}(X,Y)$$

这些量提供了 (X,Y) 的很多数字特征. 例如，由切比雪夫不等式知对任意常数 k，$|X-\mu_X|\leqslant k\sigma_X$ 发生的概率大于 $1-\frac{1}{k^2}$，又比如：如果 $|\rho_{XY}|=1$，那么，X 与 Y 几乎必然落在一条直线上. ■

将均值、方差、协方差和(严)平稳性等概念扩展到随机过程上去，就有下列定义.

定义 5.1.7 设 $\{X_t:t\in I\}$ 是一个随机过程.

(i) 如果对任意 t，均有 $EX_t^2<\infty$，则称过程 $\{X_t\}$ 为一个**二阶过程**或 $\boldsymbol{L_2}$ **过程**.

(ii) 如果对任意 t，均有 $E|X_t|<\infty$，则称 $\mu(t)=EX(t)$ 为 X 的**均值函数**.

(ⅲ) 称 $\gamma_X(s, t)=\mathrm{Cov}(X_s, X_t)(t, s\in I)$ 为 X 的**(自)协方差函数**.

180

(ⅳ) 称 $r_X(s, t)=E(X_sX_t)(s, t\in I)$ 为 X 的**相关函数**.

(ⅴ) 如存在函数 $\widetilde{\gamma}$，使得 $\gamma_X(s, t)=\widetilde{\gamma}_X(|s-t|)$，则称 $\{X_t:t\in I\}$ 是**(弱)平稳的**.

(ⅵ) 如果

$$(X_{t_1},\cdots,X_{t_n})\overset{d}{=}(X_{t_1+h},\cdots,X_{t_n+h})$$

对所有 $t_1<\cdots<t_n$，$n\geqslant 1$ 及 $h\in\mathbf{Z}$ 成立，则称 $\{X_t\}$ 是**严平稳的**.

5.2 布朗运动

5.2.1 定义及基本性质

定义 5.2.1(布朗运动，维纳过程) 如果对某个常数 $\sigma>0$，随机过程 $\{B(t)=B_t, t\in[0, \infty)\}$ 满足下列 4 个条件，则称 $\{B_t\}$ 为**布朗运动**(简记为 BM)，也称为**维纳过程**.

(ⅰ) 对任何时间点 $0\leqslant t_0<t_1<\cdots<t_n$，$n\in\mathbf{N}$，增量随机变量集

$$B(t_0),B(t_1)-B(t_0),\cdots,B(t_n)-B(t_{n-1})$$

是随机独立的.

(ⅱ) 增量服从正态分布，即存在一个常数 $\sigma>0$，对所有 t，$h\geqslant 0$，均有

$$B(t+h)-B(t)\sim N(0,\sigma^2 h).$$

等价地，对所有 $0\leqslant s\leqslant t$，有 $B(t)-B(s)\sim N(0, \sigma^2(t-s))$.

(ⅲ) 对所有 $\omega\in\Omega$，路径 $t\mapsto B(t, \omega)$ 是连续的.

(ⅳ) $B(0)=0$(布朗运动始于 0 点).

称常数 $\sigma=1$ 的布朗运动为**标准布朗运动**. 在本书中，所涉及的布朗运动通常是指标准布朗运动.

如果给定了一个布朗运动 $\{B_t\}$，它可以产生一个自然滤子 $\mathcal{F}_t=\sigma\{B_s: 0\leqslant s\leqslant t\}$，$t\geqslant 0$，使得 $\{B_t\}$ 是 $\{\mathcal{F}_t\}$ 适应的. 但是，人们还想考虑一个更大的滤子，需要它包含比布朗运动的路径更多的信息. 例如，

$$\mathcal{F}_t=\sigma\{B_s,X_s:0\leqslant s\leqslant t\},t\geqslant 0,$$

其中 $\{X_t\}$ 是另一个定义在同一概率空间上的过程. 这就引出下列更一般的定义.

定义 5.2.2 如果下列条件成立，

(ⅰ) B_t 是 $\mathcal{F}_t$ 适应的；

(ⅱ) 对所有 $s\leqslant t$，增量 B_t-B_s 独立于 σ 代数 $\mathcal{F}_s$；

(ⅲ) 存在某个常数 $\sigma>0$，对所有 $s\leqslant t$，$B_t-B_s\overset{d}{=}B_{t-s}-B_0\sim N(0, \sigma^2(t-s))$；

181

则称 $\{B_t\}$ 为一个**关于滤子 $\mathcal{F}_t$ 的布朗运动或维纳过程**.

特别地，如果 $\sigma=1$，则称 $\{B_t\}$ 为一个关于滤子 $\mathcal{F}_t$ 的标准布朗运动.

由定义可看出，在布朗运动中，B_t 及它的增量 $B_t-B_s(s\leqslant t)$ 都服从正态分布. 在下面还将看到，增量之间的独立性意味着，布朗运动在 $k\in\mathbf{N}$ 个固定时间点 t_1，$\cdots$，t_k 的样本

对应的随机向量(B_{t_1}，…，B_{t_k})是多维正态随机向量.

定义 5.2.3(高斯过程) 如果对任意 t_1，…，t_k，$k\in\mathbf{N}$，取自随机过程$\{X_t,\ t\in I\}$的样本随机向量(X_{t_1}，…，X_{t_k})满足

$$(X_{t_1},\cdots,X_{t_k}) \sim N(\mu_{t_1,\cdots,t_k},\Sigma_{t_1,\cdots,t_k})$$

其中，

$$\mu=\mu_{t_1,\cdots,t_k}=E(X_{t_1},\cdots,X_{t_n})'$$

$$\Sigma=\Sigma_{t_1,\cdots,t_k}=(\mathrm{Cov}(X_{t_i},\ X_{t_j}))_{1\leqslant i,j\leqslant k}.$$

则称随机过程$\{X_t,\ t\in I\}$为**高斯过程**.

如果 $\Sigma_{t_1,\cdots,t_k}$是可逆的，那么(X_{t_1}，…，X_{t_k})的密度函数为

$$f(x_1,\cdots,x_k)=\frac{1}{(2\pi)^{n/2}\sqrt{\det(\Sigma)}}\exp\left\{-\frac{1}{2}(x-\mu)'\Sigma^{-1}(x-\mu)\right\},$$

其中，$x=(x_1,\ \cdots,\ x_k)\in\mathbf{R}^k$.

▶**例 5.2.4** 一个布朗运动是一个高斯过程，且均值函数为

$$m(t)=EB(t)=0,\quad t\geqslant 0.$$

对任意 $0\leqslant s\leqslant t$，由 $B(s)$与 $B(t)-B(s)$的独立性知，自协方差函数为

$$\begin{aligned}\gamma(s,t)&=\mathrm{Cov}(B(s),B(t))\\&=\mathrm{Cov}(B(s),B(s)+[B(t)-B(s)])\\&=\mathrm{Cov}(B(s),B(s))\\&=\sigma^2 s,\end{aligned}$$

因此，

$$\gamma(s,t)=\sigma^2(s\wedge t)=\sigma^2\min(s,t).$$ ◀

现考虑取自一个标准布朗运动的 n 个固定时间点 $t_1<\cdots<t_n$ 的样本随机向量

$$(B_{t_1},\cdots,B_{t_n}),\tag{5.1}$$

因为它是随机向量(B_{t_1})，$B_{t_2}-B_{t_1}$，…，$B_{t_n}-B_{t_{n-1}}$)的线性变换，而且根据布朗运动的性质，(B_{t_1}，$B_{t_2}-B_{t_1}$，…，$B_{t_n}-B_{t_{n-1}}$)的分量是独立的，且每个分量均是正态的，所以(B_{t_1}，…，B_{t_n})也是正态的.

182

现讨论(B_{t_1}，…，B_{t_n})的 n 元密度函数. 先考虑 $n=2$ 的特殊情形并计算在 $B_s=x$ 的条件下 B_{s+t}的条件分布. 首先介绍转移密度的概念：设$\{X_t\}$是在集合$\mathcal{X}\subset\mathbf{R}$ 上取值的随机过程，如果存在某个可测函数 $\varphi(x,\ y)$，使得

$$P(X_{s+t}\leqslant z\mid X_s=x)=\int_{-\infty}^{z}\varphi(x,y)\mathrm{d}y,\quad z\in\mathbf{R},x\in\mathcal{X},$$

则称 φ 为$\{X_t\}$的**转移密度**. 更一般地，对固定的 $t_1<\cdots<t_n$，考虑

$$\begin{aligned}&P(B_{t_n}\leqslant x_n\mid B_{t_i}=x_i,i=0,\cdots,n-1)\\&=P(B_{t_n}-B_{t_{n-1}}\leqslant x_n-B_{t_{n-1}}\mid B_{t_i}=x_i,i=0,\cdots,n-1)\\&=P(B_{t_n}-B_{t_{n-1}}\leqslant x_n-x_{n-1})\end{aligned}$$

$$= \int_{-\infty}^{x_n - x_{n-1}} \frac{1}{\sqrt{2\pi(t_n - t_{n-1})}} \exp\left(-\frac{u^2}{2(t_n - t_{n-1})}\right) du.$$

设 $x_n = z$，$t_n = s+t$，$t_{n-1} = s$ 和 $n=2$，则有

$$P(B_{s+t} \leqslant z | B_s = x) = \int_{-\infty}^{z-x} \frac{1}{\sqrt{2\pi t}} \exp\left(-\frac{u^2}{2t}\right) du$$

$$\underset{y=x+u}{=} \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi t}} \exp\left(-\frac{(x-y)^2}{2t}\right) dy.$$

因此，对所有 $t>0$ 及 $x \in \mathbf{R}$，函数

$$y \mapsto p(t,x,y) := \varphi_{(0,t)}(x-y) = \frac{1}{\sqrt{2\pi t}} \exp\left(-\frac{(x-y)^2}{2t}\right), \quad y \in \mathbf{R},$$

是给定 $B_s = x$ 的条件下 B_{s+t} 的密度函数，其中 $\varphi_{(\mu,\sigma^2)}$ 表示均值为 μ，方差为 $\sigma^2 > 0$ 的正态分布的密度函数. 由于 $B_s \sim \varphi_{(0,s)}$，则

$$(x,y) \mapsto \varphi_{(0,s)}(x) \varphi_{(0,t)}(x-y)$$

是随机向量(B_s，B_{s+t})的密度函数，其中(B_s，B_{s+t})是从形如 $\pi_t(B) = B_t$ 的两个**投影**获得的随机向量. 将这种联合密度的乘积形式推广到有限个投影的情形，就得到布朗运动的**有限维分布**(fidis). 在讨论它之前，先给出随机向量(5.1)(记其为 X)是正态分布的一个简洁证明.

首先注意到

183

$$\Sigma = \mathrm{Cov}(X) = \mathrm{diag}(t_1, t_2 - t_1, \cdots, t_n - t_{n-1}).$$

设

$$A = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ 1 & 1 & 0 & \cdots & 0 \\ \vdots & & & & \vdots \\ 1 & 1 & 1 & \cdots & 1 \end{pmatrix}.$$

则

$$(B_{t_1}, \cdots, B_{t_n})' = AX \sim N(0, \Sigma^*),$$

其中，

$$\Sigma^* = A\Sigma A' = \begin{pmatrix} t_1 & t_1 & \cdots & \cdots & \cdots & t_1 \\ t_1 & t_2 - t_1 & t_2 & \cdots & \cdots & t_2 \\ t_1 & t_2 & t_3 - t_2 & t_3 & \cdots & t_3 \\ \vdots & \vdots & & \ddots & & \vdots \\ \vdots & \vdots & & & \ddots & \vdots \\ t_1 & t_2 & \cdots & \cdots & \cdots & t_n - t_{n-1} \end{pmatrix}.$$

引理 5.2.5(有限维分布)　固定 $0 = t_0 < t_1 < \cdots < t_n$，$n \in \mathbf{N}$. 设$\{B_t\}$是一个布朗运动(始于 0 点)，那么，对任意$(x_1, \cdots, x_n) \in \mathbf{R}^n$，$(B_{t_1}, \cdots, B_{t_n})$的密度函数可以按下列方式得到：

$$(x_1,\cdots,x_n)\mapsto\varphi_n(x_1,\cdots,x_n)=\prod_{i=1}^{n}\varphi_{(0,\sigma(t_i-t_{i-1}))}(x_i-x_{i-1}),$$

其中，$x_0=0$.

证明 用数学归纳法证明. 在整个证明中都假定 $\sigma=1$. 当 $n=1$ 时，由 $B_{t_1}\sim N(0,\ t_1-t_0)$ 知结论成立. 假设关于 n 结论成立，现证明关于 $n+1$ 结论也成立. 在条件 $(B_{t_1},\ \cdots,\ B_{t_n})=(x_1,\ x_2,\ \cdots,\ x_n)$ 下，对任意 Borel 集 $A_1,\ \cdots,\ A_{n+1}$，概率 $P(B_{t_1}\in A_1,\ \cdots,\ B_{t_{n+1}}\in A_{n+1})$ 等于

$$\int_{A_1}\cdots\int_{A_{n+1}}P(B_{t_{n+1}}\in A_{n+1}\,|\,(B_{t_1},\cdots,B_{t_n})=(x_1,\cdots,x_n))\mathrm{d}P_{(B_{t_1},\cdots,B_{t_n})}(x_1,\cdots,x_n).$$

回忆下列事实：如果 $(B_{t_1},\ \cdots,\ B_{t_n})$ 的密度函数为 $\varphi_n(x_1,\ x_2,\ \cdots,\ x_n)$，那么

$$P(B_{t_1}\in A_1,\cdots,B_{t_n}\in A_n)=\int_{A_1}\cdots\int_{A_n}\varphi_n(x_1,\cdots,x_n)\mathrm{d}x_1\cdots\mathrm{d}x_n$$

且对任意 $\mathrm{d}P_{(B_{t_1},\cdots,B_{t_n})}$ 可测函数 $h(x_1,\ x_2,\ \cdots,\ x_n)$，期望

184

$$\begin{aligned}Eh(B_{t_1},\cdots,B_{t_n})&=\int\cdots\int h(x_1,\cdots,x_n)\mathrm{d}P_{(B_{t_1},\cdots,B_{t_n})}(x_1,\cdots,x_n)\\&=\int\cdots\int h(x_1,\cdots,x_n)\varphi_n(x_1,\cdots,x_n)\mathrm{d}x_1\cdots\mathrm{d}x_n.\end{aligned}$$

现取 $A_i=(-\infty,\ x_i]$，$i=1,\ \cdots,\ n+1$，则有

$$\begin{aligned}&P(B_{t_n+1}\leqslant x_{n+1}\,|\,(B_{t_1},\cdots,B_{t_n})=(x_1,\cdots,x_n))\\&\quad=P(B_{t_{n+1}}-B_{t_n}\leqslant x_{n+1}-x_n\,|\,B_{t_i}=x_i,i=1,\cdots,n)\\&\quad=\int_{-\infty}^{x_{n+1}-x_n}\varphi(0,t_{n+1}-t_n)(z)\mathrm{d}z=\int_{-\infty}^{x_{n+1}}\varphi_{(0,t_{n+1}-t_n)}(u-x_n)\mathrm{d}u.\end{aligned}$$

从而有

$$\begin{aligned}&P(B_{t_1}\leqslant x_1,\cdots,B_{t_n}\leqslant x_n,B_{t_{n+1}}\leqslant x_{n+1})\\&\quad=\int_{-\infty}^{x_1}\cdots\int_{-\infty}^{x_n}\left\{\int_{-\infty}^{x_{n+1}}\varphi_{(0,t_{n+1}-t_n)}(u-x_n)\mathrm{d}u\right\}\mathrm{d}P_{(B_{t_1},\cdots,B_{t_n})}(x_1,\cdots,x_n)\\&\quad=\int_{-\infty}^{x_1}\cdots\int_{-\infty}^{x_n}\int_{-\infty}^{x_{n+1}}\varphi_{(0,t_{n+1}-t_n)}(u-x_n)\mathrm{d}u\prod_{i=1}^{n}\varphi_{(0,t_i-t_{i-1})}(x_i-x_{i-1})\mathrm{d}x_1\cdots\mathrm{d}x_n\\&\quad=\int_{-\infty}^{x_1}\cdots\int_{-\infty}^{x_{n+1}}\prod_{i=1}^{n+1}\varphi_{(0,t_i-t_{i-1})}(x_i-x_{i-1})\mathrm{d}x_1\cdots\mathrm{d}x_{n+1},\end{aligned}$$

这就证明了引理的结论. ■

由引理 5.2.5 知：如果一个随机事件只与布朗运动在有限个时间点的值有关，那么就能够计算这个随机事件的概率. 这自然使人联想到对依赖于整个路径的事件，上述结论仍成立吗？事实上，可以证明，这也是成立的，因为有限维分布确定了整个过程的分布.

定义 5.2.6(布朗桥) 设 $\{B_t,\ t\in[0,\ 1]\}$ 是一个标准的布朗运动，称过程

$$B^0(t)=B_t-tB(1),\quad t\in[0,1]$$

为一个**布朗桥**.

显然，对任意 t 有 $E(B_t^0)=0$，直接计算易知，对任意 $0\leqslant s\leqslant t\leqslant 1$，

$$\mathrm{Cov}(B_t,B_s)=s(1-t)$$

成立，且 B_t^0 是高斯过程. 对固定的 $s\in(0,1)$，由

$$E(B_s^0 B_1) = E(B_s B_1 - sB_1^2) = \min(s,1) - s = 0$$

知 B_s^0 与 B_1 相互独立. 类似地，对任意 $0<s_1<\cdots<s_n<1$，由 B^0 在这些时间点上的投影得到的随机向量$(B_{s_1}^0, \cdots, B_{s_n}^0)$是独立于 B_1 的，从而$\{B_s^0, s\in(0,1)\}$独立于 B_1.

185

布朗运动本身及很多由它导出的过程都是鞅. 下列命题列举了与布朗运动有关的三个最重要的鞅(关于自然滤子$\mathcal{F}_t=\sigma(B_s: s\leqslant t)$).

命题 5.2.7 设$\{B_s: t\geqslant 0\}$是一个标准的布朗运动(关于$\{\mathcal{F}_t\}$)，那么，下列过程是$\mathcal{F}_t$ 鞅.

(i) $\{B_t: t\geqslant 0\}$.

(ii) $\{B_t^2 - t: t\geqslant 0\}$.

(iii) $\left\{\exp\left(\sigma B_t - \frac{\sigma^2}{2}t\right): t\geqslant 0\right\}$，$\sigma>0$.

证明 首先验证(i)，设 $s\leqslant t$，注意到 $E(B_t - B_s)=0$. 因为 B_s 是$\mathcal{F}_s$ 可测的，且增量 $B_t - B_s$ 是独立于$\mathcal{F}_s$ 的，则有

$$E(B_t \mid \mathcal{F}_s) = B_s + E(B_t - B_s \mid \mathcal{F}_s) = B_s + E(B_t - B_s) = B_s.$$

下面验证(ii)，注意到，

$$\begin{aligned}
E(B_t^2 \mid \mathcal{F}_s) &= E((B_s + B_t - B_s)^2 \mid \mathcal{F}_s) \\
&= B_s^2 + 2B_s \underbrace{E(B_t - B_s \mid \mathcal{F}_s)}_{=0} + E((B_t - B_s)^2 \mid \mathcal{F}_s) \\
&= B_s^2 + \underbrace{E((B_t - B_s)^2 \mid \mathcal{F}_s)}_{=t-s} \\
&= B_s^2 + (t-s),
\end{aligned}$$

从而 $E(B_t^2 - t \mid \mathcal{F}_s) = B_s^2 - s$. 最后验证(iii)，因为对任意 $s\leqslant t$ 有

$$\begin{aligned}
E(e^{\sigma B_t} \mid \mathcal{F}_s) &= E(e^{\sigma(B_s + B_t - B_s)} \mid \mathcal{F}_s) \\
&= e^{\sigma B_s} \cdot E(e^{\sigma(B_t - B_s)} \mid \mathcal{F}_s) \\
&= e^{\sigma B_s} E(e^{\sigma(B_t - B_s)}).
\end{aligned}$$

下面计算 $E(e^{\sigma(B_t - B_s)}) = E(e^{\widetilde{Z}})$，其中 $\widetilde{Z} \overset{d}{=} \sigma(B_t - B_s) \sim N(0, \sigma^2(t-s))$. 首先计算 $Z\sim N(0,1)$, $b\in\mathbf{R}$ 时的 $E(e^{bZ})$：

$$\begin{aligned}
E(e^{bZ}) &= \int e^{bx} \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx \\
&= \int \frac{1}{\sqrt{2\pi}} \exp\left(\frac{-(x-b)^2 + b^2}{2}\right) dx \\
&= \exp\left(\frac{b^2}{2}\right) \underbrace{\int \frac{1}{\sqrt{2\pi}} \exp\left(\frac{-(x-b)^2}{2}\right) dx}_{=1} \\
&= \exp\left(\frac{b^2}{2}\right).
\end{aligned}$$

186

现设 $b=\sigma\sqrt{t-s}$，从而得到

$$E(e^{\sigma B_t}\mid\mathcal{F}_s)=e^{\sigma B_s}e^{\frac{\sigma^2(t-s)}{2}}.$$

因此有

$$E\left(e^{\sigma B_t-\frac{\sigma^2}{2}t}\mid\mathcal{F}_s\right)=e^{\sigma B_s-\frac{\sigma^2}{2}s}.$$

■

出现在命题 5.2.7 中的过程是非常重要的.

定义 5.2.8(几何布朗运动) 称随机过程

$$X_t(\omega)=X_0\exp(\mu t+\sigma B_t(\omega)),\quad t>0,\omega\in\Omega,$$

为带漂移率 $\mu\in\mathbf{R}$，波动率 $\sigma>0$ 的**几何布朗运动**，其中，B_t 是一个标准的布朗运动.

推论 5.2.9 一个带漂移率 $\mu=-\frac{\sigma^2}{2}$，波动率 $\sigma>0$ 的几何布朗运动是一个$\mathcal{F}_t$ 鞅.

对于一个给定的布朗运动$\{B_t,\ t\in[0,\ 1]\}$，前面介绍的有限维分布的结果描述了它的概率情况，如果将过程中的时间变为 $t\in[0,\ \lambda]$，即考虑过程 $B_{\lambda t}$，$t\in[0,\ 1]$，结果会如何呢?

引理 5.2.10(缩放性质) 设 $B=\{B_t\}$是一个标准的布朗运动(始于 0 点)，$\lambda>0$ 是一个固定的常数，那么

$$\{B_{\lambda t}:t\geqslant 0\}\overset{d}{=}\{\lambda^{1/2}B_t:t\geqslant 0\}.$$

更一般地，对 $t_1<\cdots<t_n$，$n\in\mathbf{N}$，随机向量$(B_{\lambda t_1},\ \cdots,\ B_{\lambda t_n})$与 $\lambda^{1/2}(B_{t_1},\ \cdots,\ B_{t_n})$同分布.特别地，$\{\lambda^{-1/2}B_{\lambda t}:t\geqslant 0\}$仍是一个标准布朗运动.

定义 5.2.11 设$\{X_t\}$是一个取值在集合$\mathcal{X}\subset\mathbf{R}$ 上的$\mathcal{F}_t$ 适应过程，如果对任一可测函数 $f:\mathcal{X}\to\mathbf{R}$，均存在可测函数 $g:\mathcal{X}\to\mathbf{R}$，使得

$$E(f(X_t)\mid\mathcal{F}_s)=g(X_s)$$

对所有 $0\leqslant s\leqslant t\leqslant T$ 成立，则称$\{X_t\}$为一个关于$\mathcal{F}_t$ 的**马尔可夫过程**.

定理 5.2.12 一个布朗运动$\{B_t,\ t\geqslant 0\}$是一个关于$\mathcal{F}_t=\sigma(B_s:s\leqslant t)$，$t\geqslant 0$ 的马尔可夫过程.

证明 首先注意到，对给定的$\mathcal{F}_s$，B_s 的值是确定的，设 $B_s=x$，则对所有 $\omega\in\Omega$，有

$$E(f(B_t)\mid\mathcal{F}_s)(\omega)=E(f(x+(B_t-B_s))\mid\mathcal{F}_s)(\omega).$$

187

因为 B_t-B_s 是独立于$\mathcal{F}_s$ 的，所以

$$E(f(x+B_t-B_s))=\int f(x+y)\varphi_{(0,\sigma(t-s))}(y)\mathrm{d}y=g(x)$$

关于 x 的可测函数. ■

5.2.2 布朗运动与中心极限定理

设$\{B_t\}$是定义在概率空间$(\Omega,\ \mathcal{F},\ P)$上的布朗运动，定义随机变量 $\xi_{ni}(\omega):(\Omega,\ \mathcal{F},\ P)\to(\mathbf{R},\ \mathcal{B})$为

$$\xi_{ni}(\omega)=\sqrt{n}\cdot\left(B\left(\frac{i}{n},\omega\right)-B\left(\frac{i-1}{n},\omega\right)\right),\quad i=1,\cdots,n.$$

则 ξ_{ni} 具有下列性质:

（ⅰ）对任意 $n\in\mathbf{N}$，ξ_{n1}，ξ_{n2}，…，ξ_{nn}是独立的.

（ⅱ）$\xi_{ni}\sim N(0,\sigma^2)$，$1\leqslant i\leqslant n$，$n\in\mathbf{N}$.

且由定义知

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{k}\xi_{ni}=B_{k/n},$$

则有

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{\lfloor nt\rfloor}\xi_{ni}=B_{\lfloor nt\rfloor/n}\underset{n\to\infty}{\to}B_t\sim N(0,\sigma^2 t).$$

如果撇开 ξ_{ni} 的具体定义不谈，那么我们注意到性质（ⅰ）和性质（ⅱ）意味着 ξ_{ni}，$1\leqslant i\leqslant n$，$n\geqslant 1$，构成随机变量的三角形阵列，其中每行随机变量都是独立同分布的．由中心极限定理知，对任一固定的 t，$\lfloor nt\rfloor^{-1/2}\sum_{i=1}^{\lfloor nt\rfloor}\xi_{ni}$ 依分布收敛于 $N(0,\sigma^2)$，从而当 $n\to\infty$时有

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{\lfloor nt\rfloor}\xi_{ni}\xrightarrow{d}B_t\sim N(0,\sigma^2 t).\tag{5.2}$$

由上面的观察可以看出，在更一般的假设条件下，布朗运动应该为上述形式的部分和提供一种正确的渐近极限，而且是在分布收敛意义下(有限维分布，甚至更广泛意义下)的极限．将上面的和式概括为下列定义：

定义 5.2.13 设$\{X_{ni}:i=1,\cdots,n.,n\geqslant 1\}$是一个随机变量阵列，称过程

$$S_n(t)=\frac{1}{\sqrt{n}}\sum_{i=1}^{\lfloor nt\rfloor}X_{ni},\quad t\in[0,1],$$

188 为一个部分和过程.

在将上面的观察结果应用于二叉树模型中的股票价格的收敛性(使用它可以得到著名的 Black-Scholes 公式)之前，有必要讨论一下中心极限定理的一个明显的推广：一个部分和过程的渐近分布是布朗运动．依分布收敛概念的适当推广是弱收敛的概念，设$\{X, X_n\}$是取值于可分的完备度量空间$\{S, d\}$(d 为度量)上的随机过程序列，如果

$$\int f(X_n)\mathrm{d}P\to\int f(X)\mathrm{d}P,\quad n\to\infty$$

对任一有界函数 $f:S\mapsto\mathbf{R}$(它关于度量 d 是连续的)成立，则称 X_n 当 $n\to\infty$时弱收敛于 X，记为

$$X_n\Rightarrow X,$$

如果 $S=\mathbf{R}$，且 $d(x)=|x|$，$x\in\mathbf{R}$，则弱收敛就是依分布收敛．由著名的 Donsker 不变原理知：当$\{S_n\}$的增量是均值为 0，方差为 1 的 i. i. d. 时，有

$$S_n\Rightarrow B,\quad n\to\infty.$$

从而在由右连左极函数组成的一类非常广泛的集合 A 上，$P(S_n\in A)$可由 $P(B\in A)$近似．该结论也可由下列事实得到：对任一连续映射 f，$P(f(S_n)\in A)$收敛于 $P(f(B)\in A)$．有关弱收敛的内容在这里不作进一步的讨论了，感兴趣的读者可阅读附录 B. 1.

理所当然地，现在可以对股价过程的中心极限定理(它是推导 Black-Scholes 定价公式的基础)做更深入的了解了.

重温定理 4.7.3

首先回顾一下定理的结论，设 P_n 表示 n 期二叉树模型的等价鞅测度，则在 $\{P_n\}$ 下，$\{S_n(t)\}$ 依分布收敛于随机变量

$$S_t = S_0\exp\left(\left(r-\frac{\sigma^2}{2}\right)t+\sigma B_t\right), \quad t>0.$$

其中，$\{B_t\}$ 是概率空间 $(\Omega, \mathcal{F}, P^*)$ 上的标准布朗运动，即在概率测度 P^* 下，布朗运动 B 的漂移率为 0，波动率为 1，且 P^* 是 P_n 由 $\lim\limits_{n\to\infty} p_n=1/2$ 确定的极限. 由此得到贴现股价

$$S_n^*(t) = e^{-rt}S_n(t)$$

依分布收敛于一个几何布朗运动，即当 $n\to\infty$ 时，

$$S_n^*(t) \xrightarrow{d} S_t^* = S_0\exp(-t\sigma^2/2+\sigma B),$$

由推论 5.2.9 知，在 P^* 下，上面的极限过程是一个鞅. 使用 Skorohod 空间中弱收敛的相关结论，可以将定理 4.7.3 中的依分布收敛推广为：当二叉树模型中 $n\to\infty$，(贴现)部分和价格过程 $\{S_n^*(t):t\in[0, T]\}$ 弱收敛于几何布朗运动 $\{S_t^*:t\in[0, T]\}$，因此价格过程近似于一个连续时间的过程.

189

上面的结果表明几何布朗运动应该是建立股票价格的理想数学连续时间模型. 且存在一个概率测度，使得在该测度下，贴现股票价格过程是一个鞅.

5.2.3 路径性质

定义 5.2.14 设 $\{X_t\}$ 和 $\{Y_t\}$ 是同一概率空间 $(\Omega, \mathcal{F}, P)$ 上的两个随机过程，如果

$$P(X_t\neq Y_t)=0, \forall t\in I,$$

则称 $\{Y_t\}$ 是 $\{X_t\}$ 的一个**修正**，也称 $\{X_t\}$ 和 $\{Y_t\}$ 是**等价过程**.

引理 5.2.15 如果 $\{Y_t\}$ 是 $\{X_t\}$ 的一个修正，那么它们是分布相等的，即它们的有限维(边缘)分布相同.

证明 如能证明下列更强的结论

$$P(X_{t_1}=Y_{t_1},\cdots,X_{t_n}=Y_{t_n})=1, \forall 0\leqslant t_1,\cdots,t_n\leqslant T, n\in\mathbf{N}. \qquad (*)$$

则引理得证. 事实上，如果(*)式不成立，则有

$$P(X_{t_1}=Y_{t_1},\cdots,X_{t_n}=Y_{t_n})<1.$$

则存在集合 $\Omega_0\in\mathcal{F}$，满足

$$\forall\omega\in\Omega_0:\exists j(\omega)\in\{1,\cdots,n\}:X_{t_{j(\omega)}}(\omega)\neq Y_{t_{j(\omega)}}(\omega).$$

记 $A_i=\{X_{t_i}\neq Y_{t_i}\}$，则 $\Omega_0=\bigcup\limits_{i=1}^{n}A_i$，且至少存在一个 i 使得 $P(A_i)>0$，这是一个矛盾. ■

▶**例 5.2.16** 本例将表明可以改变一个过程的值，使得改变前后的过程是等价的. 例如：可以改变布朗运动在一个不影响它分布的随机时间点的值. 设 $\{B_t\}$ 是一个布朗运动，$U\sim\exp(1)$，对 $t\geqslant 0$，令

$$V_t=\begin{cases}B_t, & U\neq t,\\ B_t+1, & U=t,\end{cases}$$

显然$\{V_t\}$的路径不再连续. 要证明$\{V_t\}\overset{d}{=}\{B_t\}$，由引理5.2.15，只需证明$\{V_t\}$是$\{B_t\}$的一个修正即可. 事实上，对所有$t\geqslant 0$，

$$P(V_t \neq B_t) = P(U = t) = 0,$$

这表现：尽管$\{\omega$:对所有$t\in[0,\infty)$，$V_t(\omega)=B_t(\omega)\}=\varnothing$从而

$$P(V_t = B_t : \forall t \in [0,\infty)) = 0,$$

即这两个过程没有一条路径是相同的，但它们仍是分布相同的. ◀

190

上例表明：等价过程不一定有相同的路径.

定义 5.2.17　如果

$$P(X_t = Y_t, \quad \forall t \in I) = 1,$$

则称$\{X_t\}$与$\{Y_t\}$是**无区别的**.

5.2.4 多维布朗运动

设$B(t)$是一个标准布朗运动，$B_1(t)$，…，$B_n(t)$是$B(t)$的n个相互独立的复制，称过程

$$t \mapsto (B_1(t),\cdots,B_n(t))', \quad t \geqslant 0,$$

为路径在欧几里得空间R^n中的n维布朗运动. 由构造知，该过程是坐标独立的. 现考虑线性组合

$$X_i(t) = a_{i1}B_1(t) + \cdots + a_{in}B_n(t), \quad i = 1,\cdots,m,$$

其中，$a_{ij}\in\mathbf{R}$，则得到一列高斯过程. 如果令

$$\boldsymbol{A} = \begin{bmatrix} \boldsymbol{a}_1' \\ \vdots \\ \boldsymbol{a}_m' \end{bmatrix},$$

其中，$\boldsymbol{a}_i'=(a_{i1}, \cdots, a_{in})$，$i=1$，…，$m$. 则有

$$X(t) = (X_1(t),\cdots,X_n(t))' = \boldsymbol{A}B(t).$$

特别地，由B_1，…，B_n的独立性知

$$\begin{aligned}
\mathrm{Cov}(X_i(s),X_j(t)) &= \mathrm{Cov}(a_{i1}B_1(s) + \cdots + a_{in}B_n(s), a_{j1}B_1(t) + \cdots a_{jn}B_n(t)) \\
&= a_{i1}a_{j1}\mathrm{Cov}(B_1(s),B_1(t)) + \cdots + a_{in}a_{jn}\mathrm{Cov}(B_n(s),B_n(t)) \\
&= (a_{i1}a_{j1} + \cdots + a_{in}a_{jn})\min(s,t) \\
&= \boldsymbol{a}_i'\boldsymbol{a}_j\min(s,t).
\end{aligned}$$

如果用$\mathrm{Cov}(X(s), X(t))$表示协方差的$m\times m$矩阵，则有

$$\mathrm{Cov}(X(s),X(t)) = \min(s,t)\boldsymbol{AA}'.$$

n维布朗运动的一般定义如下.

定义 5.2.18　设$\{B_i(t), t\geqslant 0\}$，$i=1$，…，n.，是n个1维随机过程，称n维随机过程$t\mapsto\boldsymbol{B}(t)=(B_1(t), \cdots, B_n(t)\}$为**协方差矩阵为$\boldsymbol{\Sigma}$的$n$维布朗运动**，如果下列条件成立：

(ⅰ) 对所有$t\geqslant 0$，$E(\boldsymbol{B}(t))=0$；

191

(ⅱ) $\{B(t)\}$的有限维分布是高斯过程；

(ⅲ) 协方差满足：

$$\mathrm{Cov}(B_i(s), B_j(t)) = E(B_i(s)B_j(t)) = \sigma_{ij}\min(s,t)$$

对所有 $1 \leqslant i$, $j \leqslant n$，及 s, $t \geqslant 0$ 成立，其中，σ_{ij} 是 Σ 的元素.

5.3 连续性与可微性

下面的著名定理为判断一个随机过程的路径光滑性提供了一个简便的矩准则.

定理 5.3.1(Kolmogorov-Chentsov) 设 $\{X_t, t>0\}$ 是一个实值过程，如果它满足：对任意 $T>0$，均存在常数 α, β, $C>0$，使得下式成立

$$E|X_s - X_t|^{\alpha} \leqslant C \cdot |t-s|^{1+\beta}, \forall s,t \in [0,T],$$

那么，存在$\{X_t\}$的一个修正，使得对任意 $\gamma \in (0, \beta/\alpha)$，该修正是路径局部 γ 阶 Hölder-连续的.

现将 Chentsov 准则应用于标准布朗运动 B_t，由

$$B_t - B_s \sim N(0, |t-s|) \stackrel{d}{=} \sqrt{|t-s|} \cdot U, \quad U - N(0,1),$$

知

$$E|B_t - B_s|^{2n} = |t-s|^n E|U|^n.$$

取 $n=2$，则有 $E|B_t - B_s|^4 = |t-s|^2$，由 Chentsov 准则知存在一个具有 Hölder-连续路径的布朗运动．记 $\alpha=2n$，$\beta=n-1$，则

$$\frac{\beta}{\alpha} = \frac{n-1}{2n} \uparrow \frac{1}{2}.$$

因此，对任意 $0<\gamma<1/2$，布朗运动是 γ 阶 Hölder-连续的.

前面已经阐明了布朗运动有连续的样本路径，从而有 $\lim\limits_{s\to t} B(s) = B(t)$（对其中一个修正）.

定义 5.3.2

(ⅰ) 取值于 $\mathbf{R}$ 中的随机过程$\{X_t\}$称为**概率连续的**，如果它满足：

$$P\left(\lim_{s\to t} X_s = X_t\right) = 1.$$

(ⅱ) 取值于 $\mathbf{R}$ 中的随机 L_2-过程$\{X_t\}$称为**均方连续的(m.s. 连续的)**，如果它满足：对所有 t，

$$\lim_{s\to t} X_s = X_t \quad \text{在 } L_2 \text{ 中成立},$$

192

即

$$\lim_{s\to t} E|X_s - X_t|^2 = 0.$$

定义 5.3.2 能够容易地推广到取值于 $\mathbf{R}^d$ 中的随机过程上去．事实上，只需将定义 5.3.2 中的 $|\cdot|$ 换成向量范数就可以了.

对一个随机过程，可以通过考察其相关函数 $r_X(s, t) = E(X_s X_t)$ 来判断其是否均方连续.

引理 5.3.3 一个过程$\{X_t\}$是均方连续的，当且仅当

$$当(u,v)\to(t,t)时,\quad r_X(u,v)\to r_X(t,t).\tag{5.3}$$

特别地，如果$\{X_t\}$是平稳过程，且$r_X(t)=r_X(0, t)$在$t=0$点连续，则$\{X_t\}$是均方连续的.

证明 先证充分性. 因为当$s\to t$时，有$(s, t)\to(t, t)$，由(5.3)式知

$$r_X(s,t)\to r_X(t,t),$$

充分性由下列关系式马上得到

$$E(X_s-X_t)^2=r_X(s,s)-2r_X(s,t)+r_X(t,t),$$

下面证明必要性. 因为

$$|r_X(u,v)-r_X(t,t)|=|E(X_uX_v)-E(X_tX_t)|$$

不大于

$$|E(X_uX_v)-E(X_tX_u)|+|E(X_tX_u)-E(X_tX_t)|,$$

对上式的第一项使用柯西-施瓦兹不等式得到

$$|E(X_uX_v)-E(X_tX_u)|=|E(X_u[X_v-X_t])|\leqslant\sqrt{E(X_u^2)E(X_v-X_t)^2},$$

当$(u, v)\to(t, t)$时，它收敛于 0，同理可证第二项也收敛于 0，从而得到(5.3)式成立. ■

5.4 自相似与分数布朗运动

设B_t，$t\geqslant 0$，是一个标准布朗运动，由比例性质(引理 5.2.10)知，对任意$\lambda>0$

$$\{B_{\lambda t}:t\geqslant 0\}\overset{d}{=}\{\lambda^{1/2}B_t:t\geqslant 0\}.\tag{5.4}$$

这意味着：时间比例的变化$t\mapsto\lambda t$与状态空间的变化$x\mapsto\lambda^{1/2}t$是等效的. 一个过程如果满足时间比例的变化与状态空间的某种变化(在分布意义下)是等效的，则称这个过程是自相似的. 精确定义如下.

定义 5.4.1(自相似) 如果过程$\{X_t:t\geqslant 0\}$满足：对任意$\lambda>0$，均存在$\upsilon=\upsilon(\lambda)>0$，使得

193

$$\{X_{\lambda t}:t\geqslant 0\}\overset{d}{=}\{\upsilon X_t:t\geqslant 0\}.$$

则称该过程为**自相似的**. 当$\upsilon=\lambda^H$时，称H为 **Hurst 指数**，称$\{X_t:t\geqslant 0\}$是 **Hurst 指数为 H 的自相似过程**，称$D=1/H$为$\{X_t\}$的**(统计)分数维数**.

布朗运动的自相似性可由它的协方差函数推出. 事实上，因为

$$E(B_{\lambda s},B_{\lambda t})=\min(\lambda s,\lambda t)=\lambda\min(s,t)=\lambda E(B_s,B_t)=E(\lambda^{1/2}B_s)(\lambda^{1/2}B_t),$$

将这一结论推广到有限维分布中去就得到(5.4)式成立.

自然会想到这样一个问题：是否存在协方差函数$c(s, t)$，使得

$$c(\lambda s,\lambda t)=\lambda^H c(s,t)$$

对某个$H\neq 1/2$成立呢？回答是肯定的：考虑

$$c_H(s,t)=\frac{1}{2}(|s|^{2H}+|t|^{2H}-|s-t|^{2H}),s,t\in\mathbf{R}.\tag{5.5}$$

显然，对 $0<H\leqslant 1$，函数 $c_H(s, t)$是非负定的，且满足

$$c_H(\lambda s,\lambda t) = \lambda^{2H} c_H(s,t), \quad s,t \in \mathbf{R}.$$

因为对任一非负定函数，必存在一个零均值的高斯过程，使其协方差函数为这个非负定函数. 设$\{X_t, t\geqslant 0\}$是协方差函数为 $c_H(s, t)$的高斯过程，类似前面布朗运动的讨论有：

$$E(X_{\lambda s}X_{\lambda t}) = c_H(\lambda s,\lambda t) = \lambda^{2H}E(X_s X_t) = E(\lambda^H X_s)(\lambda^H X_t),$$

从而得到

$$\{X_{\lambda t}:t \geqslant 0\} \stackrel{d}{=} \{\lambda^H X_t:t \geqslant 0\}.$$

即$\{X_t:t\geqslant 0\}$是 Hurst 指数为 H 的自相似过程. 那么$\{X_t:t\geqslant 0\}$存在连续的修正吗？下面验证 Kolmogorov-Chentsov 准则的条件：对所有 $0\leqslant s, t$，有

$$\begin{aligned} E|X_t - X_s|^2 &= E(X_t^2) - 2E(X_t X_s) + E(X_s^2) \\ &= c_H(t,t) + c_H(s,s) - 2c_H(s,t) = |t-s|^{2H}. \end{aligned}$$

因此，当 $H>1/2$ 时，Kolmogorov-Chentsov 准则的条件满足，其中对应的 $\alpha=2$，$\beta=2H-1$. 下面讨论 $H<1/2$ 的情形，首先注意到：$\{X_t\}$有零均值的平稳增量，且增量的方差为

$$E(B_t^H - B_s^H)^2 = |t-s|^{2H},$$

即

$$B_t^H - B_s^H \sim N(0,|t-s|^{2H}).$$

这意味着增量的矩为

$$E(B_t^H - B_s^H)^{2k} = \frac{2k!}{k!2^k}|t-s|^{2Hk}.$$

对给定的 $0<H\leqslant 1/2$，选取 k 满足 $2Hk>1$，则 Kolmogorov-Chentsov 准则的条件满足.

194

定义 5.4.2(分数布朗运动) 一个均值为零，协方差函数为(5.5)的连续高斯过程称为**一个标准分数布朗运动**.

注意，当 $H=1/2$ 时，分数布朗运动就是一个标准布朗运动.

一个 Hurst 指数为 $0<H\leqslant 1$ 的分数布朗运动 B_t^H，$t\geqslant 0$，具有比例性质

$$\{B_{\lambda t}^H:t \geqslant 0\} \stackrel{d}{=} \{\lambda^H B_t^H:t \geqslant 0\}.$$

将分数布朗运动的性质概括起来就是：一个 Hurst 指数为 $0<H\leqslant 1$ 的分数布朗运动是一个始于 0，具有平稳的正态增量，且是 Hurst 指数为 H 的自相似过程的过程.

由分数布朗运动具有平稳，独立的差可以导出下列定义：

定义 5.4.3(分数高斯噪声) 设$\{B_t^H, t\geqslant 0\}$是一个 Hurst 指数为 $0<H<1$ 的分数布朗运动，称序列

$$\xi_t = B_{t+1}^H - B_t^H, \quad t = 0,1,\cdots$$

为一个标准分数高斯噪声.

5.5 计数过程

5.5.1 泊松过程

设 τ_1，τ_2，…是随机时刻，例如，τ_i 可能是由于一个突发事件而导致股价急变的时间点．可以使用不同的方法将这些随机时刻用一个函数表示出来．例如，使用一个示性过程

$$I_t = \sum_{i=1}^{\infty} \mathbf{1}_{\{\tau_i\}}(t), \quad t \geqslant 0.$$

那么，对任意 $i \in \mathbf{N}$，当且仅当 $t=\tau_i$ 时，$I_t=1$．在实际中，常常需要累计突发事件的总数，因此下面讨论由这些随机时刻导出的**计数过程**或**点过程**

$$N_t = \sum_{i=1}^{\infty} \mathbf{1}(\tau_j \leqslant t), \quad t \geqslant 0.$$

如果 $\sup_n \tau_n = \infty$，则称 N_t 为**没有爆破的计数过程**．显然 N_t 是到 t 时刻为止发生的突发事件数．注意到：N_t 是一个右连左极函数．更精确地说，N_t 是一个右连续的分段常函数，在随机时刻 τ_i 的跳跃为 1．

易证，当且仅当所有 τ_i 均为停时的时候，N_t 是适应过程．

注意，如果 $\tau_1 \leqslant \tau_2 \leqslant \cdots$，则有

$$\tau_i \leqslant t \Leftrightarrow N_t \geqslant i, \quad \tau_i \leqslant t < \tau_{i+1} \Leftrightarrow N_t = u.$$

195

定义 5.5.1 设$\{N_t : t \geqslant 0\}$是一个计数过程，如果它是适应的，没有爆破，有独立平稳增量，且服从下列分布

$$N_t - N_0 \sim Poi(\lambda t), \quad t \geqslant 0,$$

其中 $\lambda \in (0, \infty)$是一个常数，则称$\{N_t, t \geqslant 0\}$是一个**泊松过程**，λ 称为强度参数．其对应的中心化过程

$$N_t^c = N_t - \lambda t, \quad t \geqslant 0$$

称为一个**补偿泊松过程**．

计数过程在$[0, t]$内有 k 个跳点的概率为

$$P(N_t = k) = \frac{(\lambda t)^k}{k!} \mathrm{e}^{-\lambda t}, k = 0, 1, \cdots,$$

从而得到 N_t 的均值、方差、协方差函数分别为

$$E(N_t) = \lambda t, \quad \mathrm{Var}(N_t) = \lambda t, \quad \mathrm{Cov}(N_s, N_t) = \min(s, t)\lambda,$$

其中，s，$t \geqslant 0$．容易看到，$E(N_t)/t = \lambda$ 是跳跃的平均数目（单位时间内的跳跃数），称之为**平均到达率**．

下列引理是显而易见的．

引理 5.5.2 N_t^c 是 0 均值的鞅．

因为服从参数为 λ 的泊松随机变量的特征函数为 $u \mapsto \exp(\lambda[\mathrm{e}^{\mathrm{i}u} - 1])$，从而，泊松过程的特征函数 $\varphi_{N_t}(u) = E(\mathrm{e}^{u\mathrm{i}N_t})$ 为

$$\varphi_{N_t}(u)=\exp(\lambda t[e^{iu}-1]), u\in\mathbf{R},$$

对补偿泊松过程，其特征函数为

$$\varphi_{N_t^c}(u)=E(e^{iu(N_t-\lambda t)})=\exp(\lambda t[e^{iu}-1-iu]), u\in\mathbf{R}.$$

5.5.2 复合泊松过程

在上一小节介绍的随机时刻的例子中，市场有影响股价的不可预测的信息，假定发生在时刻 τ_i 的第 i 个事件引起的股价波动为 Y_i，因为到时刻 t 为止，N_t 个事件分别引起的跳跃为 Y_1，…，Y_{N_t}，所以，到时刻 t 为止，所有突发事件对股价的总影响为

$$X_t=\sum_{i=1}^{N_t}Y_i.$$

196

定义 5.5.3 设 $\{N_t:t\geqslant 0\}$ 是一个泊松过程，$\{Y_n:n\in\mathbf{N}\}$ 是一个 i.i.d. 随机变量序列，且 $\{N_t\}$ 与 $\{Y_n\}$ 独立，则称 $X_t=\sum_{i=1}^{N_t}Y_i, t\geqslant 0$ 为一个**复合泊松过程**. 它的中心化过程

$$X_t^c=X_t-\lambda tE(Y_1),\quad t\geqslant 0$$

称为**补偿泊松过程**.

如果 $Y_1\in L_2$，那么一个复合泊松过程的均值、方差、协方差函数分别为

$$E(X_t)=\lambda tE(Y_1),\quad \mathrm{Var}(X_t)=\lambda tE(Y_1^2),\quad \mathrm{Cov}(X_s,X_u)=\lambda\min(s,u)E(Y_1^2).$$

显然，一个复合泊松过程的样本路径是不连续的，但是，下列结果表明：当 $s\approx t$ 时，增量 X_t-X_s 大于 $\varepsilon>0$ 的概率是很小的，且当 $s\to t$ 时，这个概率趋于 0.

引理 5.5.4 一个复合泊松过程是依概率连续的.

证明 固定 $\varepsilon>0$，如果没有跳跃，那么

$$|X_t-X_s|=\left|\sum_{i=N_s+1}^{N_t}Y_i\right|=0,$$

即

$$\{\omega:N_t(\omega)=N_s(\omega)\}\subset\{\omega:|X_t(\omega)-X_s(\omega)|=0\},$$

所以 $\{|X_t-X_s|>0\}\subset\{N_t>N_s\}$. 从而当 $s\to t$ 时有

$$P(|X_t-X_s|\geqslant\varepsilon)\leqslant P(N_t-N_s>0)=1-\exp(-\lambda(t-s))\to 0.$$ ■

引理 5.5.5 一个复合泊松过程 $X_t=\sum_{i=1}^{N_t}Y_i(t\geqslant 0)$ 具有平稳增量，且增量与过去时间的增量是相互独立的.

证明 因为当 $s<t$ 时，

$$X_t-X_s=\sum_{i=N_s+1}^{N_t}Y_i=\sum_{i=1}^{N_t-N_s}Y_{N_s+i}\overset{d}{=}\sum_{i=1}^{N_t-N_s}Y_i.$$

由 $N_t-N_s\overset{d}{=}N_{t-s}$ 知

$$X_t-X_s\overset{d}{=}\sum_{i=1}^{N_{t-s}}Y_i=X_{t-s},$$

197

增量的平稳性得证. 下面证明对 $0\leqslant r<s<t$, X_s-X_r 与 X_t-X_s 是独立的. 对所有 Borel 集 A 与 B, 类似上面的讨论有

$$P(X_s-X_r\in A,X_t-X_s\in B)=P\Big(\sum_{i=1}^{N_s-N_r}Y_i\in A,\sum_{i=1}^{N_t-N_s}Y_i'\in B\Big),$$

其中, Y_1', Y_2', …是与 Y_1 同分布的 i. i. d. 随机变量序列, 且独立于 Y_1, Y_2, …与$\{N_t\}$. 由 N_s-N_r 与 N_t-N_s 是独立的, 且均独立于$\{Y_i, Y_i'\}$, 可得

$$P\Big(\sum_{i=1}^{N_s-N_r}Y_i\in A,\sum_{i=1}^{N_t-N_s}Y_j'\in B\Big)=P\Big(\sum_{i=1}^{N_s-N_r}Y_i\in A\Big)P\Big(\sum_{i=1}^{N_t-N_s}Y_i'\in B\Big).$$

因此, X_s-X_r 与 X_t-X_s 的独立性得证.

下面计算复合泊松过程的特征函数 $\varphi_{X_t}(u)=E(\mathrm{e}^{\mathrm{i}uX_t})$. 设 φ_Y 表示 Y_1 的特征函数, 则在条件 $N_t=k$ 下, 对 $u\in\mathbf{R}$ 有

$$\begin{aligned}\varphi_{X_t}(u)&=E\Big(E\Big(\mathrm{e}^{\mathrm{i}u\sum_{j=1}^{N_t}Y_j}\,|\,N_t\Big)\Big)\\&=\sum_{k=0}^{\infty}E\Big(\mathrm{e}^{\mathrm{i}u\sum_{j=1}^{k}Y_j}\Big)P(N_t=k)\\&=\mathrm{e}^{-\lambda t}\sum_{k=0}^{\infty}\frac{(\varphi_Y(u)\lambda t)^k}{k!}\\&=\exp(\lambda t[\varphi_Y(u)-1]),\end{aligned}$$

由这个公式, 立即得到补偿复合泊松过程的特征函数为

$$\varphi_{X_t^c}(u)=E(\mathrm{e}^{\mathrm{i}u(X_t-\lambda tE(Y_1))})=\exp(\lambda t[\varphi_Y(u)-1-\mathrm{i}uE(Y_1)]).$$

为了对介绍 Lévy 过程做准备, 下面用带跳跃过程的分布函数 F 来表示这个公式. 由
$\varphi_Y(u)=\int_{-\infty}^{\infty}\mathrm{e}^{\mathrm{i}us}\mathrm{d}F(s),1=\int_{-\infty}^{\infty}\mathrm{d}F(s)$ 及 $E(Y_1)=\int_{-\infty}^{\infty}s\mathrm{d}F(s)$,
得到

$$\varphi_{X_t^c}(u)=\exp\Big(\lambda t\int_{-\infty}^{\infty}(\mathrm{e}^{\mathrm{i}us}-1-\mathrm{i}us)\mathrm{d}F(s)\Big),u\in\mathbf{R}.$$

如果 Y_j 服从均值 $\mu\in\mathbf{R}$, 方差 $\sigma^2>0$ 的正态分布, 那么, 复合泊松过程的密函数 f_{X_t} 可以具体表示出来. 事实上, 由

198

$$P\Big(\sum_{j=1}^{k}Y_j\leqslant x\,|\,N_t=k\Big)=P\Big(\sum_{j=1}^{k}Y_j\leqslant x\Big)=\Phi_{(k\mu,k\sigma^2)}(x),$$

知

$$\begin{aligned}P(X_t\leqslant x)&=\sum_{k=0}^{\infty}P(X_t\leqslant x\,|\,N_t=k)P(N_t=k)\\&=\sum_{k=0}^{\infty}\Phi_{(k\mu,k\sigma^2)}(x)\frac{(\lambda t)^k}{k!}\mathrm{e}^{-\lambda t},\end{aligned}$$

从而得到密度函数

$$f_{X_t}(x) = \sum_{k=0}^{\infty} \frac{1}{\sqrt{2\pi k\sigma^2}} \exp\left(-\frac{1}{2}\frac{(x-k\mu)^2}{k\sigma^2}\right)\frac{(\lambda t)^k}{k!}\mathrm{e}^{-\lambda t}, \quad x \in \mathbf{R}.$$

5.6 Lévy 过程

前面已经学习了带漂移的布朗运动，它可以表示成 $\mu t+\sigma B_t$，$t\geqslant 0$，其中，$\mu\in\mathbf{R}$，$\sigma>0$，$\{B_t\}$是一个标准的布朗运动. 带漂移的布朗运动是一类具有下列性质的过程:(ⅰ)有连续路径；(ⅱ)具有独立平稳增量. 为了推广到带跳跃的过程，下面讨论将条件(ⅰ)放松为依概率连续.

定义 5.6.1 设$\{X_t\}$是一个适应的随机过程，$X_0=0$(a. s.)，且满足下列条件

(ⅰ) $\{X_t\}$有平稳(齐次)增量，且增量独立于过去的增量，即

$$X_{t+s}-X_t \text{ 独立于} \mathcal{F}_t \text{ 且 } X_{t+s}-X_t \overset{d}{=} X_s, \text{对任意 } s\geqslant 0.$$

(ⅱ) $\{X_t\}$是依概率连续的.

则称$\{X_t\}$是一个 **Lévy 过程.**

下面是一个不常用的定义:如$\{X_t\}$是一个满足 $X_0=0$(a. s.)，有独立平稳增量，且是依概率连续的，则称该过程为**本性 Lévy 过程**.

可以证明:任一 Lévy 过程必存在一个右连左极的修正. 因此，通常将右连左极加入到 Lévy 过程的定义中. 将 Lévy 过程的定义推广到 d 维过程是直接的. 任一 d 维 Lévy 过程必存在一个修正，使得每一个坐标分量过程有连续路径.

易证:Lévy 过程的和仍是 Lévy 过程.

引理 5.6.2 如果 $X_t(t\geqslant 0)Y_t(t\geqslant 0)$是两个 Lévy 过程，那么 $X_t+Y_t(t\geqslant 0)$仍是 Lévy 过程.

由这个引理知:一个带漂移的布朗运动加上一个复合泊松过程仍是一个 Lévy 过程，即

$$X_t = \mu t + \sigma B_t + \sum_{i=1}^{N_t} Y_i, \quad t\geqslant 0,$$

是一个 Lévy 过程，其中:$\{N_t:t\geqslant 0\}$及$\{Y_n:n\in\mathbf{N}\}$独立于$\{B_t:t\geqslant 0\}$. 显然，X_t 是一个右连左极过程.

199

由于 Lévy 过程$\{X_t\}$有独立平稳增量，因而其边际分布 P_{X_t}是无限可分的. 所以其特征函数 $\varphi(u)=E(\mathrm{e}^{\mathrm{i}uX_t})$，$u\in\mathbf{R}$ 满足

$$\varphi_{t+s}(u) = \varphi_t(u)\varphi_s(u),$$

从而得到

$$\varphi_t(u) = \exp(t\psi(u)),$$

其中，$\psi(\theta)$是累积量母函数. 由 Lévy-Khintchine 公式(1.6)得到带三元组(b, C, ν)的下列公式

$$\varphi_t(u) = \exp\left\{\mathrm{i}ub_t - \frac{1}{2}uC_t\theta + \int_{\mathbf{R}^d}(\mathrm{e}^{\mathrm{i}ux} - 1 - \mathrm{i}ux\mathbf{1}(|x| \leqslant 1))\mathrm{d}\nu(x)\right\},$$

将上面两个公式结合起来，则有

$$B_t = tb, \quad C_t = tC, \quad \nu_t(x) = t\nu(x),$$

$$\psi(u) = \mathrm{i}\theta' b - \frac{1}{2}uC\theta + \int_{\mathbf{R}^d}(\mathrm{e}^{\mathrm{i}ux} - 1 - \mathrm{i}ux\mathbf{1}(|x| \leqslant 1))\nu(\mathrm{d}x).$$

回忆一下补偿复合泊松过程的特征函数公式

$$\varphi_{X_t^c}(u) = \exp\left(\lambda t\int_{-\infty}^{\infty}(\mathrm{e}^{\mathrm{i}us} - 1 - \mathrm{i}us)\mathrm{d}F(s)\right), u \in \mathbf{R}.$$

设 $\mathrm{d}\nu(s) = \dfrac{\mathrm{d}F(s)}{\lambda}$，则可以看出：$X_t^c$是一个 Lévy 过程.

最后，讨论具有无限 Lévy 测度的 Lévy 过程的一个标准结构. 设 $N_t^{(k)}(k \geqslant 1)$是一列独立的泊松过程，参数分别为 $\lambda_k > 0$，$k \in \mathbf{N}$，又设 $\mu_k(k \geqslant 1)$是一列实数. 那么，对任一 $N \in \mathbf{N}$，

$$X_t^{(N)} = \sum_{k=1}^{N}\mu_k(N_t^{(k)} - \lambda_k t)$$

是一个 Lévy 过程，其 Lévy 测度(读者自己验证)为

$$\nu^{(N)} = \sum_{k=1}^{N}\lambda_k\delta_{\{\mu_k\}},$$

其特征函数为

$$\varphi_t^{(N)}(u) = \exp\left\{t\int_{-\infty}^{\infty}(\mathrm{e}^{\mathrm{i}ux} - 1 - \mathrm{i}ux)\nu^{(N)}(\mathrm{d}x)\right\}.$$

且 $X_t^{(N)}$ 的 L_2 极限存在，极限为

200

$$X_t = \sum_{k=1}^{\infty}\mu_k(N_t^{(k)} - \lambda_k t).$$

如果假定

$$\sum_{k=1}^{\infty}\lambda_k\mu_k^2 < \infty,$$

那么极限过程也是一个 Lévy 过程，Lévy 测度为

$$\nu(x) = \sum_{k=1}^{\infty}\lambda_k\delta_{\{\mu k\}}(x),$$

且满足 $\nu(\mathbf{R}) = \Sigma_k{}^{\lambda_k}$，它可能是无限的.

5.7 评注与延伸阅读

最早涉及布朗运动的论文是 Bachelier (1900)，Einstein(1905)及 Wiener(1923). 介绍布朗运动基本知识的著作有 Wiersema(2008)和 Mikosch(1998). 系统地介绍布朗运动的概念与性质的著作可以参考 Grigoriu(2002)和 Shiryaev(1999). 建立在高等概率方法之上的

深入讨论布朗运动的著作有 Durrett(1996)，Karatzas 与 Shreve(1991)，Revuz 与 Yor (1999). 全面介绍布朗运动与经济联系的著作有 Davidson(1994).

参考文献

Bachelier L. (1900) Théorie de la spéculation. *Annales Scientifiques de l' É.N.S.* 3ᵉ **17**, 21–86.

Davidson J. (1994) *Stochastic Limit Theory: An Introduction for Econometricians*. Advanced Texts in Econometrics. The Clarendon Press Oxford University Press, New York.

Durrett R. (1996) *Probability: Theory and Examples*. 2nd edn. Duxbury Press, Belmont, CA.

Einstein A. (1905) On the movement of small particles suspended in a stationary liquid demanded by the molecular-kinetic theory of heat. *Ann. Phys.* **17**, 549–560.

Grigoriu M. (2002) *Stochastic Calculus: Applications in Science and Engineering*. Birkhäuser Boston Inc., Boston, MA.

Karatzas I. and Shreve S.E. (1991) *Brownian Motion and Stochastic Calculus*. vol. 113 of *Graduate Texts in Mathematics* 2nd edn. Springer-Verlag, New York.

Mikosch T. (1998) *Elementary Stochastic Calculus—with Finance in View*. vol. 6 of *Advanced Series on Statistical Science & Applied Probability*. World Scientific Publishing Co. Inc., River Edge, NJ.

Revuz D. and Yor M. (1999) *Continuous Martingales and Brownian Motion*. vol. 293 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]* 3rd edn. Springer-Verlag, Berlin.

Shiryaev A.N. (1999) *Essentials of Stochastic Finance: Facts, Models, Theory*. vol. 3 of *Advanced Series on Statistical Science & Applied Probability*. World Scientific Publishing Co. Inc., River Edge, NJ. Translated from the Russian manuscript by N. Kruzhilin.

Wiener N. (1923) Differential spaces. *J. Math. Phys.* **2**, 131–174.

Wiersema U. (2008) *Brownian Motion Calculus*. John Wiley & Sons Inc., New York.

201

第6章　Itô积分

本章介绍随机积分与Itô积分的基本知识．首先，本章将以一种适当的方式把第4章(见4.2节公式(4.1))中刻画自融资交易策略中的价值过程

$$I_t = \int \varphi_r \mathrm{d}S_r = \sum_{r=1}^{t} \varphi_r (S_r - S_{r-1}) \tag{6.1}$$

这种离散形式的鞅变换推广到 S_t 为布朗运动的连续时间上去，它不能通过对每一个 ω 使用 Riemann-Stieltjes 积分来实现．Kiyoshi Itô推广了 Wiener Norbert 的工作，巧妙地将这种积分定义为由公式(6.1)给出的 Riemann-Stieltjes 和的均方极限．在学习完随机 Stieltjes 积分后，本章将定义被积函数为相当广泛的一类过程的 Itô积分．我们限定 Itô积分中的积分变量为布朗运动，并简要讨论了如何求更一般过程的积分．

其次，本章将引入 Itô过程类．粗略地说，Itô过程是一个 Riemann 积分 $\int_0^t \mu_s \mathrm{d}s$ 与一个 Itô积分的简单和，其中 μ_s 是随机的，但对每个 ω 它是可积的．著名的Itô公式表明：一个布朗运动的光滑函数可以表示成一个 Itô过程；更一般地，一个 Itô过程的光滑函数仍是一个 Itô过程．在本章中，将使用 Itô公式来验证诸如(广义)几何布朗运动，Ornstein-Uhlenbeck 过程等基本过程是随机微分方程的解，这些过程是构建金融模型的基础．

最后，本章还将简略地介绍遍历扩散过程，它是覆盖各种金融模型的一个重要分支．本章还讨论当扩散过程中不能求出解的精确表达式时，求数值近似解的 Euler 逼近方法，Euler 逼近方法也为对扩散过程的样本值进行统计估计建立了基础．

6.1　全变差与二次变差

(全)变差是度量一个函数局部振幅的量．设 $f:[0, T]\to\mathbf{R}$ 是一个连续可微函数，在 $[0, t_1]$ 上，它是严格递增的，在 $[t_1, t_2]$ 上，它是严格递减的，在 $[t_2, T]$ 上，它又是严格递增的，其中，$0<t_1<t_2<T$．如果 f 在 t_1 点达到极大值，在 t_2 点达到极小值，那么用数

$$\int |\mathrm{d}f| = V(f) = f(t_1) - f(0) + (f(t_1) - f(t_2)) + (f(T) - f(t_2))$$

来度量 f 的振幅是可行的．由上式可以看出，$\int |\mathrm{d}f|$ 可以表示为

$$\int |\mathrm{d}f| = \int_0^{t_1} f'(r)\mathrm{d}r + \int_{t_1}^{t_2} - f'(r)\mathrm{d}r + \int_{t_2}^{T} f'(r)\mathrm{d}r = \int_0^T |f'(r)|\mathrm{d}r.$$

一般函数的全变差定义如下．

定义 6.1.1　函数 $f:[0, T]\to\mathbf{R}$ 的**(全)变差**定义为

$$\int |\mathrm{d}f| = V(f) = \sup \sum_{i=0}^{n-1} |f(t_{i+1}) - f(t_i)|,$$

其中，上式的值可以为 $+\infty$，sup 是对 $[0, T]$ 的所有分法 $\Pi_n: 0=t_0<\cdots<t_n=T$，$n\in\mathbf{N}$ 取上确界．如果 $\int |\mathrm{d}f|<\infty$，那么称 f 是**有界变差**函数．

容易看出所有的有界变差函数构成一个向量空间，记该空间为$BV=BV([0, T])$.

评注 6.1.2

（i）可以直接将定义 6.1.1 推广到定义域为$\mathbf{R}$或$[0, +\infty)$的函数.

（ii）如果函数$f:[0, \infty)\to\mathbf{R}$满足对任意$0<t<\infty$，$f$在$[0, t]$上的限制$f|_{[0,t]}:[0, t]\to\mathbf{R}$是有界变差，那么称$f$是**局部有界变差**函数.

（iii）如果f有连续导数，那么$\int|\mathrm{d}f|=\int_0^T|f'(r)|\mathrm{d}r$.

（iv）**Hahn-Jordan 分解**：任一有界变差函数可以分解为

$$f=f^{+}-f^{-},$$

其中f^{+}，f^{-}均为非减函数.

定义 6.1.3 函数$f:[0, T]\to\mathbf{R}$的**二次变差**定义为极限

$$[f,f](t)=[f,f]_t=\lim_{\|\Pi_n\|\to 0}\sum_{i=0}^{n-1}[f(t_{i+1})-f(t_i)]^2,$$

204

其中，$\|\Pi_n\|$表示分法$\Pi_n:0=t_0<\cdots<t_n=t$的步长，即

$$\|\Pi_n\|=\max_{0\leqslant i\leqslant n-1}|t_{i+1}-t_i|.$$

如果该极限不存在，那么称二次变差不存在.

注意，在定义 6.1.3 中，分法Π_n中的点是依赖于n的，为了使表达式简单，我们在记号中省略了n.

引理 6.1.4 如果f有连续导数f'，且$\int_0^T|f'(r)|\mathrm{d}r<\infty$，那么对任意$t\in[0, T]$，$[f, f]_t=0$.

证明 对任一分法$\Pi_n:0=t_0<\cdots<t_n=t$，由中值定理知，存在点$t_i^*\in[t_i, t_{i+1}]$，使得

$$\begin{aligned}\sum_{i=0}^{n-1}(f(t_{i+1})-f(t_i))^2&=\sum_{i=0}^{n-1}|f'(t_i^*)|^2(t_{i+1}-t_i)^2\\&\leqslant\|\Pi_n\|\sum_{i=0}^{n-1}|f'(t_i^*)|^2(t_{i+1}-t_i).\end{aligned}$$

其中用到了一个明显的不等式$(t_{i+1}-t_i)^2\leqslant\|\Pi_n\|(t_{i+1}-t_i)$，因此有

$$\begin{aligned}[f,f]_t&=\lim_{\|\Pi_n\|\to 0}\sum_{i=0}^{n-1}(f(t_{i+1})-f(t_i))^2\\&\leqslant\lim_{\|\Pi_n\|\to 0}\|\Pi_n\|.\lim_{\|\Pi_n\|\to 0}\sum_{i=0}^{n-1}|f'(t_i^*)|^2(t_{i+1}-t_i)=0.\end{aligned}$$

■

定义 6.1.5 设$\{X_t:t\in[0, T]\}$是一个连续时间的随机过程，$0\leqslant t_0<t_1<\cdots<t_n\leqslant T$，称

$$[X,X]_n(t)=[X,X]_{nt}=\sum_{i=0}^{n-1}(X_{t_{i+1}}-X_{t_i})^2$$

为**样本二次变差**(关于分法$\Pi_n:\{t_i:i=0, \cdots, n\}$).

定理 6.1.6 一个布朗运动$\{B_s:s\in[0, t]\}$的样本二次变差

$$Q_{nt} = Q_n(t) = \sum_{i=0}^{n-1}(B_{t_{i+1}} - B_{t_i})^2$$

在 L_2 中收敛且几乎必然收敛于 t，即当 $n\to\infty$ 时，对$[0, t]$的任何分法$\{t_i\}$，均有

$$Q_n(t) \xrightarrow{L_2, \text{a. s.}} t$$

205 成立.

证明 下面给出 L_2 收敛的一个证明. 注意到 $E(Q_{nt}-t)^2 = E(Q_{nt}-EQ_{nt})^2 + (EQ_{nt}-t)^2$，所以，只需证明当 $n\to\infty$时，$EQ_{nt}\to t$，且 $\mathrm{Var}(Q_{nt})\to 0$ 即可. 因为 Q_n 的被加项是独立的随机变量，且 $E(B_{t_{i+1}}-B_{t_i})^2 = \mathrm{Var}(B_{t_{i+1}}-B_{t_i}) = t_{i+1}-t_i$，所以

$$E(Q_{nt}) = \sum_{i=0}^{n-1}(t_{i+1}-t_i) = t.$$

为了计算 Q_{nt}的方差，回忆下列事实:如果随机变量$U\sim N(0, \sigma^2)$，那么 $EU^4 = 3\sigma^4$，因此有

$$\begin{aligned}\mathrm{Var}((B_{t_{i+1}}-B_{t_i})^2) &= E((B_{t_{i+1}}-B_{t_i})^2-(t_{i+1}-t_i))^2 \\ &= E(B_{t_{i+1}}-B_{t_i})^4 - 2(t_{i+1}-t_i)E(B_{t_{i+1}}-B_{t_i})^2 + (t_{i+1}-t_i)^2 \\ &= 3(t_{i+1}-t_i)^2 - 2(t_{i+1}-t_i)^2 + (t_{i+1}-t_i)^2 \\ &= 2(t_{i+1}-t_i)^2.\end{aligned}$$

从而得到

$$\mathrm{Var}(Q_{nt}) = \sum_{i=0}^{n-1}\mathrm{Var}((B_{t_{i+1}}-B_{t_i})^2) \leqslant 2\|\Pi_n\|\sum_{i=0}^{n-1}(t_{i+1}-t_i),$$

所以，当 $n\to\infty$时，$\mathrm{Var}(Q_{nt})\leqslant 2\|\Pi_n\|t\to 0$.

回忆 Riemann 积分 $\int_a^b f(g(t))\mathrm{d}t$ 存在的一个充分性条件:g 是可积的，f 是$[a, b]$上的一致连续函数. 因为布朗运动是几乎必然连续的，从而是几乎必然可积的，即随机变量

$$Z(\omega) = \int_a^b f(B(t,\omega))\mathrm{d}t, \quad \omega\in\Omega$$

几乎必然有定义，所以当 $n\to\infty$时，对任一满足$\|\Pi_n\|\to 0$ 的分法 $\Pi_n:\{t_i\}$(a. s.)有

$$\sum_{i=0}^{n-1}f(B_{t_i})(t_{i+1}-t_i) \to \int_a^b f(B_t)\mathrm{d}t$$

成立. 下一结果表明:在非常一般的条件下，将上一公式中的$(t_{i+1}-t_i)$换成$(B_{t_{i+1}}-B_{t_i})^2$(其期望为 $t_{i+1}-t_i$)，则它是 L_2 收敛的. ■

定理 6.1.7 设 f 是一致连续函数，且满足 $\int_0^t Ef^2(B_s)\mathrm{d}s < \infty$，则当 $n\to\infty$时，对$[0, t]$的任一满足$\|\Pi_n\|\to 0$ 的分法 $\Pi_n:\{t_i\}$，有

$$\sum_{i=0}^{n-1}f(B_{t_i})(B_{t_{i+1}}-B_{t_i})^2 \xrightarrow{L_2} \int_0^t f(B_s)\mathrm{d}s$$

成立.　206

证明　设

$$A_n = E\Big(\sum_{i=0}^{n-1} f(B_{t_i})(B_{t_{i+1}} - B_{t_i})^2 - \int_0^t f(B_s)\mathrm{d}s\Big)$$

$$B_n = \mathrm{Var}\Big(\sum_{i=0}^{n-1} f(B_{t_i})(B_{t_{i+1}} - B_{t_i})^2\Big).$$

下面证明当 $n\to\infty$ 时，$A_n\to 0$，$B_n\to 0$. 首先考虑 A_n，因为 $A_n=A_{n1}+A_{n2}$，其中：

$$A_{n1} = E\Big(\sum_{i=0}^{n-1}[f(B_{t_i})(B_{t_{i+1}} - B_{t_i})^2 - f(B_{t_i})(t_{i+1}-t_i)]\Big)$$

$$A_{n2} = E\Big(\sum_{i=0}^{n-1} f(B_{t_i})(t_{i+1}-t_i) - \int_0^t f(B_s)\mathrm{d}s\Big).$$

由对 f 的假定知，当 $n\to\infty$ 时，$A_{n2}\to 0$. 因为

$$E((B_{t_{i+1}} - B_{t_i})^2 \mid \mathcal{F}_{t_i}) = E(B_{t_{i+1}} - B_{t_i})^2 = t_{i+1}-t_i,$$

从而，

$$A_{n1} = \sum_{i=0}^{n-1} E(E(f(B_{t_i})(B_{t_{i+1}} - B_{t_i})^2 - f(B_{t_i})(t_{i+1}-t_i) \mid \mathcal{F}_{t_i})) = 0.$$

其次考虑 B_n，记

$$D_i = B_{t_{i+1}} - B_{t_i},$$
$$V_i = D_i^2 - (t_{i+1}-t_i),$$
$$U_{ij} = f(B_{t_i})f(B_{t_j})(D_i^2 - (t_{i+1}-t_i)).$$

则有

$$\begin{aligned} B_n &= \sum_{i,j=0}^{n-1} E[f(B_{t_i})f(B_{t_j})(D_i^2 - (t_{i+1}-t_i))(D_j^2 - (t_{j+1}-t_j))] \\ &= \sum_{i,j=0}^{n-1} E(U_{ij}V_j). \end{aligned}$$

当 $i<j$ 时，因为 U_{ij} 是$\mathcal{F}_i$ 可测的，而 D_j^2 是独立于$\mathcal{F}_i$ 的，故 U_{ij} 与 V_j 是独立的. 由 $EV_j=0$ 得到上式中所有 $i<j$ 的项为 0，类似的方法得到所有 $i>j$ 的项也为 0，因此

$$B_n = \sum_{i=0}^{n-1} E[f^2(B_{t_i})(D_i^2 - (t_{i+1}-t_i))^2].$$

又因为 $f(B_{t_i})$是$\mathcal{F}_{t_i}$ 可测的，而 D_i^2 是独立于$\mathcal{F}_{t_i}$ 的，所以有　207

$$\begin{aligned} E(f^2(B_{t_i})(D_i^2 - (t_{i+1}-t_i))^2) &= E(f^2(B_{t_i})E[(D_i^2 - (t_{i+1}-t_i))^2 \mid \mathcal{F}_{t_i}]) \\ &= E(f^2(B_{t_i})E[D_i^2 - (t_{i+1}-t_i)]^2), \end{aligned}$$

其中 $E(D_i^2-(t_{i+1}-t_i))^2=2(t_{i+1}-t_i)^2$，其证明过程与定理 6.1.6 的证明相同. 因此有

$$B_n = 2\sum_{i=0}^{n-1} E(f^2(B_{t_i}))(t_{i+1}-t_i)^2 \leqslant 2\|\Pi_n\| \sum_{i=0}^{n-1} g(t_i)(t_{i+1}-t_i),$$

其中，$g(t)=Ef^2(B_t)$，$t\in[0, T]$. 注意到

$$g(t) = \int Ef^2(B_t)\mathrm{d}t = \int_{-\infty}^{+\infty} f^2(x)\varphi_{(0,t)}(x)\mathrm{d}x,$$

这里 $\varphi_{(0,t)}(x)$表示 $N(0, t)$分布的密度函数. 由假定知 $\int_0^t g(t)\mathrm{d}t < \infty$ ，因此，当 $n\to\infty$时，Riemann 和 $\sum_{i=0}^{n-1} g(t_i)(t_{i+1}-t_i)$ 收敛于 $\int_0^t f^2(B_s)\mathrm{d}s$ ，从而当 $n\to\infty$时，$B_n\to 0$，定理得证. ■

6.2 随机 Stieltjes 积分

在一个 Riemann 和 $\sum_{i=0}^{n} f(t_i^*)(t_{i+1}-t_i)$ 中，被加项是函数值分别被(广义)分布函数 $\mathrm{id}(x)=x(x\in\mathbf{R})$在区间$(t_i, t_{i+1}]$上的差或区间$(t_i, t_{i+1}]$的 Lebesgue 测度 λ 加权而得的. 按照这一思路，一个很自然的想法是将分布函数 id 换成另一个具有有限(带附号的)测度的(广义)分布函数. 回想评注 6.1.2(iv)，下列推广是有效的. 设 $f:[0, T]\to\mathbf{R}$ 是一个连续函数，$H:[0, T]\to\mathbf{R}$ 有有界变差，那么可以定义积分

$$\int_0^T f(t)\mathrm{d}H(t)$$

为所谓 Riemann-Stieljes(RS)和

$$\sum_{i=0}^{n-1} f(t_i^*)(H(t_{i+1})-H(t_i)),\quad t_i^*\in[t_i,t_{i+1}] \tag{6.2}$$

的极限.

定义 6.2.1 如果当 $n\to\infty$时，对任一满足$\|\Pi_n\|\to 0$ 的分法 Π_n，RS 和(6.2)收敛，那么称 f 是 RS 可积的，称

$$\int_0^T f\mathrm{d}H = \int_0^T f(t)\mathrm{d}H(t) = \lim_{n\to\infty}\sum_{i=0}^{n-1} f(t_i^*)(H(t_{i+1})-H(t_i))$$

208 为 f 关于 H 的 **Riemann-Stieljes(RS)积分**，称 H 为**积分变量**.

类似地，可以定义形如 $\int_a^b f(t)\mathrm{d}H(t)$ 与 $\int_a^\infty f(t)\mathrm{d}H(t)$ 的 RS 积分，其中，$0\leqslant a\leqslant b<\infty$. RS 积分有下列性质.

命题 6.2.2 设 f 关于 H 是 RS 可积的，$0\leqslant a\leqslant b<\infty$，则下列结论成立.

(i) $f\mapsto\int_a^b f\mathrm{d}H$ (f 是 RS 可积的)是一个线性映射.

(ii) $\left|\int_a^b f\mathrm{d}H\right|\leqslant\|f\|_\infty\cdot\int|\mathrm{d}H|$.

(iii) 对任意有界变差函数 H_1 和 H_2，有

$$\int_a^b f\mathrm{d}(H_1+H_2) = \int_a^b f\mathrm{d}H_1 + \int_a^b f\mathrm{d}H_2.$$

(iv) 对任意 $\alpha\in\mathbf{R}$，有 $\int_a^b f\mathrm{d}(\alpha H) = \alpha\int_a^b f\mathrm{d}H$.

(v) f 关于 H 是 RS 可积的当且仅当 H 关于 f 是 RS 可积的，并且下列分部积分公式

成立：

$$\int_a^b f(x)\mathrm{d}H(x) = f(x)H(x)\big|_a^b - \int_a^b H(x-)\mathrm{d}f(x).$$

(vi) 如果 f 是连续的，H 是可微的，且 $h=H'$，那么 RS 积分 $\int_a^b f\mathrm{d}H$ 可以化为 Riemann 积分，即

$$\int_a^b f\mathrm{d}H = \int_a^b f(x)h(x)\mathrm{d}x.$$

(vii) 如果 H 是有限测度 μ_H 的(广义)分布函数，且有离散支集 $\{x_i:i\in I\}\subset[a, b]$，其中，$I$ 是某个离散集[⊖]，那么

$$\int_a^b f\mathrm{d}H = \sum_{i\in I} f(x_i)\mu_H(\{x_i\}).$$

注意，由性质(iii)和性质(iv)知，$H\mapsto\int f\mathrm{d}H$（$H$ 是有界变差的）是一个线性映射. 性质(vii)使得统计学中的许多量可以表示成 RS 积分. 例如，n 个点 x_1，…，$x_n\in\mathbf{R}$ 的算术平均 $\overline{x}=n^{-1}\sum_{i=1}^{n}x_i$ 是一个关于经验分布函数

$$F_n(x) = \frac{1}{n}\sum_{i=1}^{n}\mathbf{1}(x_i\leqslant x), x\in\mathbf{R}.$$

的 RS 积分.

209

事实上，因为经验分布函数对应的测度 $\mathrm{d}F_n$ 满足 $\mathrm{d}F_n(\{x_i\})=\frac{1}{n}$， 从而有

$$\overline{x} = \sum_{i=1}^{n} x_i\mathrm{d}F_n(\{x_i\}) = \int x\mathrm{d}F_n(x).$$

因此，RS 积分(关于经验分布函数)在统计学中有着非常重要的作用.

如果对布朗运动 B 关于有界变差函数 H 求积分，结果会怎么样呢？在 RS 积分的定义中，可以将 f 换成某个随机过程，只要它的 RS 和在某种随机意义下(比如点态意义或均方 L_2 意义下)的极限存在. 含有布朗运动的积分是非常重要的.

定理 6.2.3 设 $\{X(t):t\in[a, b]\}$ 是一个路径 a.s. 连续的高斯过程，记

$$K(s,t) = \mathrm{Cov}(X(s),X(t)), a\leqslant s,t\leqslant b.$$

设 H 是一个在所有子区间 $(c, \mathrm{d})\subset(a, b)$ 上均为有界变差的函数，如果

$$\sigma^2 = \int_a^b\int_a^b K(s,t)\mathrm{d}H(s)\mathrm{d}H(t)\in(0,\infty),$$

那么随机变量

$$\int_a^b X\mathrm{d}H = \int_a^b X(t)\mathrm{d}H(t)$$

a.s. 存在，并满足

⊖ 称集合 I 为一个离散集，如果它满足对所有 $x\in I$，x 均为孤立点，即 $\forall x\in I$，存在 x 的一个邻域 U_x，使得 $U_x\cap I=\{x\}$.

$$\int_a^b X\mathrm{d}H \sim N(0,\sigma^2).$$

证明 不失一般性，设 $a=0$，$b=1$. 取 $\delta\in(0,1)$，则当 $\delta\to 0$ 时，有 $\int_\delta^{1-\delta} X\mathrm{d}H \to \int_0^1 X\mathrm{d}H$. 另一方面，

210

$$\int_\delta^{1-\delta} X\mathrm{d}H = \lim_{n\to\infty}\sum_{i=\lfloor n\delta\rfloor}^{\lfloor n(1-\delta)\rfloor} X\left(\frac{i}{n}\right)\left[H\left(\frac{i}{n}\right)-H\left(\frac{i-1}{n}\right)\right].$$

上式中右边的和式服从 $N(0,\sigma_n^2)$分布，其中

$$\sigma_n^2 = \sum_{i,j=\lfloor n\delta\rfloor}^{\lfloor n(1-\delta)\rfloor} K\left(\frac{i}{n},\frac{j}{n}\right)\left[H\left(\frac{i}{n}\right)-H\left(\frac{i-1}{n}\right)\right]\left[H\left(\frac{j}{n}\right)-H\left(\frac{j-1}{n}\right)\right]$$

$$\to \sigma^2(\delta) = \int_\delta^{1-\delta}\int_\delta^{1-\delta} K(s,t)\mathrm{d}H(s)\mathrm{d}H(t)$$

当 $n\to\infty$时成立. 因为

$$N(0,\sigma_n^2) \xrightarrow[n\to\infty]{\mathrm{d}} N(0,\sigma^2(\delta)) \xrightarrow[\delta\to 0]{\mathrm{d}} N(0,\sigma^2),$$

结论得证. ■

例 6.2.4

(ⅰ) 对布朗运动 $B(t)$，$t\in[0,1]$，有

$$K(s,t) = \mathrm{Cov}(B(s),B(t)) = \min(s,t),\quad s,t\in[0,1],$$

因而

$$\sigma^2 = \int_0^1\int_0^1 K(s,t)\mathrm{d}s\mathrm{d}t = \frac{1}{3}.$$

由此得到随机变量 $Z(\omega)=\int_0^1 B(t,\omega)\mathrm{d}t(\omega\in\Omega)$ 是一个均值为 0 并且方差为 1/3 的正态随机变量.

(ⅱ) 对任一固定的 t，有

$$\int_0^t X(s)\mathrm{d}H(s) \sim \sigma B_t,$$

其中，B_t 表示标准布朗运动，

$$\sigma^2 = \frac{1}{t}\int_0^t\int_0^t K(s,t)\mathrm{d}H(s)\mathrm{d}H(t).$$ ◀

更一般地，如果 X 是均方连续的(即当 $|t-s|\to 0$ 时，$E(X(t)-X(s))^2\to 0$)，H 是有界变差的，那么积分 $\int_a^b X(t)\mathrm{d}H(t)$ 在对应的 RS 和均方收敛意义下存在. 下面再列举一些在 X 与 H 满足上述条件时 RS 积分的一些性质与法则.

211

(ⅰ) $\left\|\int_a^b X(t)\mathrm{d}H(t)\right\|_{L_2} \leqslant \sup_{t\in[a,b]}\|X(t)\|_{L_2}\int|\mathrm{d}H|$.

(ⅱ) $t\mapsto\int_a^t X(s)\,\mathrm{d}s$ 存在且在$[a,b]$上是 m. s. 可微的，并有

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_a^t X(s)\mathrm{d}s = X(t), \quad t\in[a,b].$$

(ⅲ) 如果 X'在$[a, b]$上是 m. s. 连续的，那么 $\int_a^b X'(t)\mathrm{d}t$ 存在，并等于 $X(b)-X(a)$.

(ⅳ) 设 $\mu(t)=EX(t)$，$t\in[a, b]$，如果 $\int_a^b X(t)\mathrm{d}H(t)$ 或 $\int_a^b H(t)\mathrm{d}X(t)$ 在均方意义下存在，那么积分 $\int_a^b H(t)\mathrm{d}\mu(t)$ 与 $\int_a^b H(t)\mathrm{d}\mu(t)$ 均存在，且

$$E\Big(\int_a^b H(t)\mathrm{d}X(t)\Big) = \int_a^b H(t)\mathrm{d}\mu(t), \quad E\Big(\int_a^b X(t)\mathrm{d}H(t)\Big) = \int_a^b \mu(t)\mathrm{d}H(t).$$

6.3 Itô积分

迄今为止，只要积分变量是一个有界变差函数，就可以定义一个积分. 自然会想到这样一个问题：布朗运动可以作为积分变量吗？如果 f 是一个有界变差函数，要计算 f 关于布朗运动的积分，可以使用分部积分公式(命题 6.2.2(ⅳ))去定义 $\int_a^b f(t)\mathrm{d}B(t)$，即

$$\int_a^b f(t)\mathrm{d}B(t) = f(t)B(t)\,|_a^b - \int_a^b B(t)\mathrm{d}f(t).$$

对更一般的确定性函数 $f\in L_2([a, b])$，Norbert Wiener 提出了下列方法：因为阶梯函数 $f_n(t) = f_0\mathbf{1}_{\{a\}}(t) + \sum_{i=1}^n f_i\mathbf{1}_{(t_i,t_{i+1})}(t), t\in\mathbf{R}$ 构成 $L_2([a, b])$在度量

$$\mathrm{d}(f,g) = \int_a^b f(t)g(t)\mathrm{d}t, \quad f,g\in L_2([a,b])$$

下的一个稠密子集，对这种函数，可以定义积分

$$\int_a^b f_n(t)\mathrm{d}B(t) = \sum_{i=1}^n f_i[B(t_{i+1})-B(t_i)].$$

上式的右边是一个均值为 0、方差为 $\sum_{i=1}^n f_i^2(t_{i+1}-t_i)$ 的正态分布随机变量，如果 $\int_a^b f(t)\mathrm{d}t < \infty$，$f_i=f(t_i)$，且当 $n\to\infty$时，分法的步长趋于 0，那么方差收敛于 $\int_a^b f(t)\mathrm{d}t$. 因此，$\int_a^b f(t)\mathrm{d}B(t)$ 可以定义为 $\int_a^b f_n(t)\mathrm{d}B(t)$ 的均方极限，这个均方极限是存在的，且是一个正态分布随机变量.

Kiyoshi Itô发现这种方法可以推广到一类更广泛的被积函数，特别地，这类函数包括了金融数学中用于记录自融资交易策略的(贴现)值的随机过程，即被积函数可以为 [212]

$$I_t = I_0 + \sum_{i=1}^t \varphi_i(S_i^* - S_{i-1}^*), t = 0,\cdots,T,$$

其中，$\{\varphi_t: t=0, \cdots, T\}$是可预报过程(即对任一 t，φ_t 是$\mathcal{F}_{t-1}$可测的)，$\{S_t: t=0, \cdots, T\}$是适应过程. 为了与 Itô积分的系列结果记号一致，记

$$\widetilde{\varphi}_t = \varphi_{t+1}, \quad t = 0,\cdots,T-1.$$

则 $I_t = I_0 + \sum_{i=1}^{t} \widetilde{\varphi}_{i-1}(S_i^* - S_{i-1}^*)$. 融资交易策略中，$\widetilde{\varphi}_{i-1}(=\varphi_i)$是$(i-1, i]$期间持有的股份数，$\widetilde{\varphi}_{i-1}$是$\mathcal{F}_{i-1}$可测的，虽然严格地说，$\widetilde{\varphi}_i$ 是适应过程，但仍称它为可预报过程. 牢牢记住：函数 $t \mapsto \widetilde{\varphi}_{i-1}\mathbf{1}_{(i-1,i]}(t)$在$(i-1, i]$内取值$\widetilde{\varphi}_{i-1}$，在其他点取值 0. 随机左连续阶梯函数

$$t \mapsto \widetilde{\varphi}_0(\omega)\mathbf{1}_{\{0\}}(t) + \sum_{i=1}^{t} \widetilde{\varphi}_{i-1}(\omega)\mathbf{1}_{(i-1,i]}(t), \quad t \in [0,T],$$

给出了每一时刻 $t\in[0, T]$持有的股份数.

定义 6.3.1 具有下列形式的随机过程称为**简单可预报过程**：

$$H_t(\omega) = H_0\mathbf{1}_{\{0\}}(t) + \sum_{i=0}^{n-1} H_i(\omega)\mathbf{1}_{(t_i,t_{i+1}]}(t),$$

其中，$0=t_0<\cdots<t_n=T$ 是$[0, T]$的一个分法，H_i 是$\mathcal{F}_i$ 可测的随机变量. 所有这类过程的全体记为$\mathcal{H}$.

注意，一个简单可预报过程可以对应一个交易策略. 对以这种简单可预报过程为被积函数的随机积分，其定义如下.

定义 6.3.2 设$\{H_t : t\in[0, T]\}$是一个简单可预报过程，称过程$\{I_t : t\in[0, T]\}$为 H 的 **Itô积分**，其中，

$$I_t = \sum_{i=0}^{k-1} H_i[B(t_{i+1}) - B(t_i)] + H_k[B(t) - B(t_k)], \quad t_k \leqslant t < t_{k+1},$$

记为

213

$$I(H) = \int H\mathrm{d}B = \int H(s)\mathrm{d}B(s) = \left\{\int_0^t H_s\mathrm{d}B_s : t \in [0,T]\right\}.$$

评注 6.3.3

(ⅰ) $I(H)$表示一族随机积分.

(ⅱ) 使用记号 $s\wedge t=\min\{s, t\}$，可以将 $I(H)$改写为下列紧凑形式：

$$\int_0^t H_s\mathrm{d}B_s = \sum_{i=0}^{n-1} H_i[B(t \wedge t_{i+1}) - B(t \wedge t_i)], \quad t \in [0,T].$$

(ⅲ) 由定义知，对 $t\in\{t_0, \cdots, t_n\}$，随机积分 $\int_0^t H_s\mathrm{d}B_s$ 为和 $\sum_{i=0}^{k-1} H_i[B(t_{i+1}) - B(t_i)]$, $k = 1,\cdots,n$.

下面的重要定理表明：Itô积分是一个鞅. 由评注 6.3.3(ⅲ) 的计算表达式可以看出离散时间的随机积分是一个鞅.

定理 6.3.4(鞅性与 Itô等距性) 设$\{B_t : t\in[0, T]\}$是一个标准布朗运动，$\mathcal{F}_t=\sigma(B_s : s\leqslant t)$，且$\{H_t\}$是一个简单可预报过程，那么 Itô积分 $\int_0^t H_s\mathrm{d}B_s, t \in [0,T]$ 是一个鞅，且对任意 $t\in[0, T]$，Itô等距性成立，即

$$E\left(\int_0^t H_r\mathrm{d}B_r\right)^2 = E\int_0^t H_r^2\mathrm{d}r = \int_0^t EH_r^2\mathrm{d}r.$$

证明 设$\{t_i\}$是$[0, T]$的任一分法，如果 $t_k<s<t\leqslant t_{k+1}$，即 s 与 t 均在区间$(t_k, t_{k+1}]$中，那么对简单可预报过程$\{H_t\}$，有

$$\int_0^t H_r \mathrm{d}B_r = \sum_{i=0}^{k-1} H_i (B_{t_{i+1}} - B_{t_i}) + H_k (B_t - B_{t_k}), \tag{6.3}$$

求 $\int_0^s H_r \mathrm{d}B_r$ 只需将上式右边表达式中的 B_t 替换成 B_s 即可．显然，右端的和式及 H_k，B_{t_k} 均是$\mathcal{F}_s$ 可测的，因此有

$$E\left(\int_0^t H_r \mathrm{d}B_r \,|\, \mathcal{F}_s\right) = \sum_{i=0}^{k-1} H_i (B_{t_{i+1}} - B_{t_i}) + H_k (E(B_t \,|\, \mathcal{F}_s) - B_{t_k}).$$

但 $E(B_t | \mathcal{F}_s) = B_s(\text{a.s.})$，如果 $t_{k-1} < s \leqslant t_k$，即 s 与 t 位于相邻的区间中，那么在条件$\mathcal{F}_s$ 下，(6.3)式右边表达式中的前 $k-1$ 项是确定的，只需计算(6.3)式右边其他各项的条件期望即可．因为

$$E(H_{k-1}(B_{t_k} - B_{t_{k-1}}) \,|\, \mathcal{F}_s) = H_{k-1}(E(B_{t_k} \,|\, \mathcal{F}_s) - B_{t_{k-1}}) = H_{k-1}(B_s - B_{t_{k-1}}),$$

H_k 是$\mathcal{F}_{t_{k-1}} \subset \mathcal{F}_s$ 可测的，且 $E(B_{t_k} | \mathcal{F}_s) = B_s$，a.s.，所以

$$E(H_k(B_t - B_{t_k}) \,|\, \mathcal{F}_s) = H_k(B_s - E(B_{t_k} \,|\, \mathcal{F}_s)) = 0.$$

因此对这种情况，结论也成立．类似地可以证明 s 与 t 位于任意区间中结论也成立． 214

下面将 Itô积分中的被积函数延拓到下列集合中的元素：

$$\mathcal{L} = \{Y:[0,T] \times \Omega \to \mathbf{R} \,|\, Y \text{ 是} \mathcal{F}_t \text{ 适应的和左连续的且} \|Y\|_{\mathcal{L}} < \infty\},$$

其中，范数定义为

$$\|Y\|_{\mathcal{L}} = \sqrt{\int_0^T E(Y_t^2)\mathrm{d}t}, \quad Y \in \mathcal{L}.$$

注意，因为

$$\|Y\|_{\mathcal{L}}^2 = \int_0^T \int_\Omega Y(t,\omega)^2 \mathrm{d}P(\omega)\mathrm{d}t,$$

所以$\|Y\|_{\mathcal{L}}$是 Y 关于概率测度 $\mathrm{d}P \otimes \mathrm{d}\lambda$ 的 L_2 范数，其中 λ 表示 Lebesgue 测度．

作为一个例子，下面证明标准布朗运动 $B=\{B_t:\ t \in [0,\ T]\}$属于$\mathcal{L}$．因为

$$E\left(\int_0^T B_t^2 \mathrm{d}t\right) = \int_0^T E(B_t^2)\mathrm{d}t = \int_0^T t\mathrm{d}t = \frac{T^2}{2} < \infty,$$

所以 $B \in \mathcal{L}$.

如果 $H \in \mathcal{H}$，那么由 Itô等距性知，

$$\|H\|_{\mathcal{L}}^2 = \int_0^T E(H_t^2)\mathrm{d}t = E\left(\int_0^T H_t \mathrm{d}B_t\right)^2 = \|I(H)\|_{L_2}^2,$$

其中$\|\cdot\|_{L_2}$表示$(\Omega,\ \mathcal{A},\ P)$中的 L_2 范数．也就是说，映射

$$I:\mathcal{H} \to L_2(\Omega,\mathcal{A},P), \quad H \mapsto I(H)$$

是等距的，这就是定理 6.3.4 名称的解释．下面证明$\mathcal{L}$中的过程可由简单可预报过程逼近． ■

引理 6.3.5 *$\mathcal{H}$是$\mathcal{L}$中的稠密子集，即对任一 $Y \in \mathcal{L}$，存在$\mathcal{H}$中子列$\{H_n\} \in \mathcal{H}$，使得*

$$\int_0^T E(H_n(s) - Y(s))^2 \mathrm{d}s \underset{n\to\infty}{\longrightarrow} 0.$$

证明 设 Y 是连续有界的，即存在某个常数 $C>0$，使得 $|Y| \leqslant C$，那么对任意 $t \in [0,\ T]$，$\omega \in \Omega$，当 $n \to \infty$时，阶梯函数

$$H_n(t) = Y(0)\mathbf{1}_{\{0\}}(t) + \sum_{i=0}^{n-1} Y\left(\frac{iT}{2^n}\right)\mathbf{1}_{\left(\frac{iT}{2^n},\frac{(i+1)T}{2^n}\right]}(t)$$

215 满足 $H_n(t, \omega) \to Y(t, \omega)$. 由控制收敛定理知,

$$\int_0^T \int_\Omega (H_n(t,\omega) - Y(t,\omega))^2 \mathrm{d}P(\omega)\mathrm{d}t \underset{n\to\infty}{\to} 0.$$

对一般情况，证明过程有一些技巧，在许多书上都可以找到. ■

在学习新内容之前，先介绍一个重要的Itô积分.

▶**例 6.3.6**$\left(\text{积分}\int B\mathrm{d}B\right)$　设函数 $F(t)$ 可导，且 $f(t)=F'(t)$，那么(通过变量代换 $z=F(t)$，$\mathrm{d}z/\mathrm{d}t=f(t)$)有

$$\int_0^T F(t)\mathrm{d}F(t) = \int_0^T F(t)f(t)\mathrm{d}t = \int_0^T z\mathrm{d}z = \frac{T}{2}.$$

如果将 $F(t)$ 换成布朗运动 $B(t)$，结果是什么呢？为了计算这个Itô积分，将用到引理6.3.5. 下面讨论将 $F(t)$ 换成布朗运动后等式是否仍成立. 为此，考虑满足条件 $\|\Pi_n\| \to 0$ 的分法列

$$\Pi_n: 0 = t_0 < \cdots < t_n = T, t_i = \frac{iT}{n}, i = 0, \cdots, T,$$

为了简化记号，设

$$W_i = B_{t_i},\quad i = 0, \cdots, n,$$

则有 $E(W_i)=0$，　$E(W_i^2)=t_i$，且当 $n\to\infty$ 时，左连续的简单可预报阶梯函数

$$H_n(t) = W_0\mathbf{1}_{\{0\}}(t) + \sum_{i=0}^{n-1} W_i \mathbf{1}_{(t_i,t_{i+1}]}(t), t \in [0,T]$$

满足 $\|H_n - B\|_L \to 0$. 事实上，当 $t\in[t_j, t_{j+1}]$ 时，由随机变量

216

$$\int_{t_j}^{t_{j+1}} (H_n(t) - B(t))^2 \mathrm{d}t,\quad j = 0, \cdots, n-1$$

的独立性知

$$\begin{aligned}
\Delta_n &= E\left(\int_0^T (H_n(t) - B(t))^2 \mathrm{d}t\right) \\
&= E\left(\sum_{j=0}^{n-1}\int_{t_j}^{t_{j+1}} (H_n(t) - B(t))^2 \mathrm{d}t\right) \\
&= \sum_{j=0}^{n-1}\int_{t_j}^{t_{j+1}} E(B_{t_j} - B_t)^2 \mathrm{d}t,
\end{aligned}$$

又因为 $E(B_{t_j}-B_t)^2=(t-t_j)$，所以当 $n\to\infty$ 时，有

$$\Delta_n = \sum_{j=0}^{n-1}\left(\frac{t^2}{2} - t_j t\right)\Bigg|_{t_j}^{t_{j+1}} = \frac{1}{2}\frac{T^2}{n} \to 0,$$

这就证明了当 $n\to\infty$ 时，$\|H_n - B\|_{\mathcal{L}} \to 0$.

为了计算 $\int_0^T H_n(s)\mathrm{d}B_s$，回顾一下定义6.3.2，有

$$\int_0^T H_n(s)\mathrm{d}B_s = \sum_{i=0}^{n-1} W_i(W_{i+1} - W_i) = \sum_{i=0}^{n-1} W_i W_{i+1} - \sum_{i=0}^{n-1} W_i^2.$$

由

$$W_iW_{i+1}=-\frac{(W_{i+1}-W_i)^2-W_{i+1}^2-W_i^2}{2}$$

得到

$$\int_0^T H_n(s)\mathrm{d}B_s=-\frac{1}{2}\sum_{i=0}^{n-1}(W_{i+1}-W_i)^2+\frac{1}{2}\sum_{i=0}^{n-1}(W_{i+1}^2-W_i^2).$$

因为 $W_{t_n}=B_T$，$W_0=B_0=0$，所以上式右边第二个和式等于 $W_n^2/2=B_T^2/2$. 由布朗运动的性质知 $W_{i+1}-W_i\sim N(0,\ t_{i+1}-t_i)$，从而有

$$E\Big(\sum_{i=0}^{n-1}(W_{i+1}-W_i)^2\Big)=\sum_{i=0}^{n-1}(t_{i+1}-t_i)=T. \tag{6.4}$$

进一步，如果 $U\sim N(0,\ 1)$，$V\sim\chi_1^2$，则有

$$(W_{i+1}-W_i)^2\overset{d}{=}(\sqrt{t_{i+1}-t_i}U)^2\overset{d}{=}(t_{i+1}-t_i)V,$$

217

再由 $\mathrm{Var}(V)=2$，增量 $W_{i+1}-W_i$，$i=0$，…，$n-1$ 相互独立得到

$$\begin{aligned}\mathrm{Var}\Big(\sum_{i=0}^{n-1}(W_{i+1}-W_i)^2\Big)&=2\sum_{i=0}^{n-1}(t_{i+1}-t_i)^2 \qquad (6.5)\\ &\leqslant 2\cdot\|\Pi_n\|\cdot T\to 0.\end{aligned}$$

组合(6.4)式与(6.5)式得到：当 $n\to\infty$ 时，

$$S_n=\sum_{i=0}^{n-1}(W_{i+1}-W_i)^2\xrightarrow{L_2}T.$$

将上述结果结合起来就有：当 $n\to\infty$ 时，

$$\int_0^T H_n(s)\mathrm{d}B_s=\frac{1}{2}(B_T^2-S_n)\xrightarrow{L_2}\frac{B_T^2-T}{2},$$

这意味着，Itô积分列 $\int_0^T H_n(s)\mathrm{d}B_s$ 在 L_2 中收敛于一个已知的随机变量$\frac{B_T^2-T}{2}$. 所以将积分 $\int_0^T B_s\mathrm{d}B_s$ 定义为这个 L_2 极限是有意义的，即定义

$$\int_0^T B_s\mathrm{d}B_s:=\frac{B_T^2}{2}-\frac{T}{2}.$$

因此，当 $F(t)$ 为布朗运动时，像 $\int_0^T F\mathrm{d}F=\frac{T^2}{2}$ 这样的基本结果不再成立. ◀

上面的例子为如何将Itô积分中被积函数扩展为 L 中的过程提供了一个思路：对任一 $Y\in\mathcal{L}$，取一列过程 $\{H_n\}\subset\mathcal{H}$，使得当 $n\to\infty$ 时，$\|Y-H_n\|_{\mathcal{L}}\to 0$. 显然，$\{H_n\}$是一个柯西列，即当 n，$m\to\infty$ 时，有

$$\|H_n-H_m\|_{\mathcal{L}}^2=\int_0^T E(H_n(t)-H_m(t))^2\mathrm{d}t\to 0,$$

218

由 I 的线性与等距性知：当 n，$m\to\infty$ 时，有

$$\left\|\int_0^T H_n\mathrm{d}B-\int_0^T H_m\mathrm{d}B\right\|_{L_2}^2=E(I(H_n)-I(H_m))^2$$

$$
\begin{aligned}
&= \| I(H_n) - I(H_m) \|_{L_2}^2 \\
&= \| I(H_n - H_m) \|_{L_2}^2 \\
&= \| H_n - H_m \|_{\mathcal{L}}^2 \to 0,
\end{aligned}
$$

因为 L_2 空间是完备的，所以 Itô积分列 $\int_0^T H_n(s)\mathrm{d}B_s$ 在 L_2 中收敛，即存在一个随机变量 $I^* \in L_2(\Omega, A, P)$，使得当 $n\to\infty$时，$\int_0^T H_n \mathrm{d}B \to I^*$ 在 L_2 中成立. 由于此极限是几乎必然唯一的，它就是所要寻找的随机积分 $\int_0^T Y\mathrm{d}B$. 将上述结果归纳起来，则有下列结果.

定义与定理 6.3.7 设$\{Y_t\}$是一个左连续的 F_t 适应过程，且满足

$$\int_0^T E(Y_t^2)\mathrm{d}t < \infty.$$

那么，存在一列关于测度 $\mathrm{d}P \otimes \mathrm{d}t$ 在 L_2 意义下逼近$\{Y_t\}$的简单可预报过程$\{H_n\}$，使得 $\int_0^T H_n \mathrm{d}B, n \geqslant 1$ 在 L_2 意义下收敛. 记其极限为

$$I(Y) = \int_0^T Y_t \mathrm{d}B_t,$$

称这个极限为 Y 关于布朗运动的**随机 Itô积分**.

显然，将区间$[0, T]$换成区间$[a, b]$，a，$b\in\mathbf{R}$ 后，上面的讨论是完全类似的. 下面给出一些 Itô积分 $\int_0^t Y_s \mathrm{d}B_s$ 作为一个连续时间 t 的随机过程的一些性质，为了简便，本书略去其证明，感兴趣的读者可参考相关文献.

引理 6.3.8 设$\{Y_t: t\in[a, b]\}$是一个左连续过程，且满足 $\int_a^b E(Y_t^2)\mathrm{d}t < \infty$.

(i) 如果 Z 是有界且$\mathcal{F}_a$可测的，那么 $ZY\in\mathcal{L}$且

$$\int_a^b ZY\mathrm{d}B = Z\int_a^b Y\mathrm{d}B.$$

219

(ii) $E\left(\int_a^b Y\mathrm{d}B \mid \mathcal{F}_a\right) = 0$.

(iii) $E\left(\left|\int_a^b Y\mathrm{d}B\right|^2 \mid \mathcal{F}_a\right) = E\left(\int_a^b Y_t^2 \mathrm{d}t \mid \mathcal{F}_a\right) = \int_a^b E(Y_t^2 \mid \mathcal{F}_a)\mathrm{d}t$.

证明见 Friedman(1975)中定理 2.8 和定理 2.9.

接下来，将 Doob 不等式 3.2.25 推广到连续时间的随机过程.

定理 6.3.9 设$\{X_t: t\in[0, T]\}$是一个右连续或左连续的下鞅，那么，当 $p\geqslant 1$ 时，对任意 $\lambda>0$，有

$$P\left(\sup_{0\leqslant s\leqslant t} |X_s| > \lambda\right) \leqslant \frac{E|X_t|^P}{\lambda^p};$$

当 $p>1$ 时，有

$$\left\| \sup_{0\leqslant s\leqslant t} |X_s| \right\|_p \leqslant \frac{p}{p-1} \sup_{0\leqslant s\leqslant t} \| X_s \|_p.$$

证明 只证明第一个不等式. 结果对连续情形基本上是成立的，因为对左连续或右连续的下鞅，有 $\sup_{0\leqslant s\leqslant t}|X_s|=\sup_{s\in A}|X_s|$，其中 $A\subset[0,t]$是一个可数稠密子集. 取一列递增的有限集 A_n，满足 $A=\bigcup_n A_n$，那么对任意 n，在 A_n 中取上确界就是取最大值，则有

$$P\Big(\sup_{s\in A_n}|X_s|>\lambda\Big)\leqslant\frac{\sup_{s\in A_n}E|X_s|^p}{\lambda}\leqslant\frac{\sup_{0\leqslant s\leqslant t}E|X_s|^p}{\lambda}.$$

注意到，当 $n\to\infty$时有$\{\sup_{x\in A_n}|X_s|>\lambda\}\uparrow\{\sup_{x\in A}|X_s|>\lambda\}$，所以

$$P\Big(\sup_{s\in A}|X_s|>\lambda\Big)=\lim_{n\to\infty}P\Big(\sup_{s\in A_n}|X_s|>\lambda\Big).$$

结合起来就得到第一个不等式. ■

定理 6.3.10 设 $Y\in\mathcal{L}$，那么Itô积分存在一个连续的修正.

证明 设$\{H_n\}\subset\mathcal{H}$，使得当 $n\to\infty$时，$\|H_n-Y\|_{\mathcal{L}}\to0$. 记

$$I_n(t)=\int_0^t H_n(s)\mathrm{d}B_s,n\geqslant1.$$

下面证明存在一个子列 I_{n_k}，$k\geqslant1$，它在 $t\in[0,T]$上是一致收敛的，从而这个(唯一的)极限是一个连续函数. 事实上，因为对每一个 n，$\{I_n(t),t\in[0,T]\}$均为一个鞅，由 Doob 不等式 3.2.25 知：对任意 n，$m\in\mathbf{N}$，$\varepsilon>0$ 有

$$P\Big(\sup_{t\in[0,T]}|I_n(t)-I_m(t)|>\varepsilon\Big)\leqslant\frac{E|I_n(T)-I_m(T)|^2}{\varepsilon^2},$$

220

其中，当 n，$m\to\infty$时，$E|I_n(T)-I_m(T)|^2=\|I_n(T)-I_m(T)\|_{L_2}^2\to0$ 成立. 所以可以取子列$n_k\to\infty$,使得

$$P\Big(\sup_{t\in[0,T]}|I_{n_{k+1}}-I_{n_k}|>\frac{1}{2^k}\Big)<\frac{1}{2^k}$$

成立. 由 Borel-Cantelli 引理知

$$N=\Big\{\sup_{t\in[0,T]}\Big|I_{n_{k+1}}-I_{n_k}\Big|>2^{-k}\quad\text{i. o.}\Big\}$$

是一个零集，所以，当 k 充分大时 $\sup\limits_{t\in[0,T]}|I_{n_{k+1}}-I_{n_k}|<2^{-k}$成立，这就证明了一个子列是一致收敛的.

评注 6.3.11 通常所指的 Itô积分 $\int_0^t Y_s\mathrm{d}B_s$ 总是指它的连续修正.

现将 Itô积分的一些重要性质和结果归纳如下.

定理 6.3.12 设$\{X_t\}$，$\{Y_t\}$是左连续的$\mathcal{L}$类过程，$0\leqslant a\leqslant b\leqslant c\leqslant T$，那么

(i) $\int_a^c X_t\mathrm{d}B_t=\int_a^b X_t\mathrm{d}B_t+\int_b^c X_t\mathrm{d}B_t$.

(ii) $\int_a^b(\lambda X_t+\mu Y_t)\mathrm{d}B_t=\lambda\int_a^b X_t\mathrm{d}B_t+\mu\int_a^b Y_t\mathrm{d}B_t$,对任意 $\lambda,\mu\in\mathbf{R}$.

(iii) $M_t=\int_0^t X_s\mathrm{d}B_s,t\in[0,T]$ 是一个鞅.

(iv) Itô积分是居中的,即 $E\Big(\int_a^b X_s\mathrm{d}B_s\Big)=0$,Itô积分也是等距的,即

$$E\left(\int_a^b X_s \mathrm{d}B_s\right)^2 = \int_a^b EX_s^2 \mathrm{d}s.$$

(ⅴ) $\left[\int_a^{\cdot} X_s \mathrm{d}B_s, \int_a^{\cdot} X_s \mathrm{d}B_s\right]_t = \int_a^t X_s^2 \mathrm{d}s, \quad t \in [a,b].$

(ⅵ) 如果$\{X_s\}$是连续的，那么对$[a, b]$的满足条件$\|\Pi_n\| \to 0$，$n\to\infty$的分法列

$$\Pi_n : a = t_{n0} < \cdots < t_{nN_n} = b$$

下列式子成立

221

$$\sum_{k=0}^{N_n-1} X_{t_{nk}} [B(t_{n,k+1}) - B(t_{nk})] \xrightarrow{P} \int_a^b X_s \mathrm{d}B_s, \quad n \to \infty.$$

(ⅶ) 对任意 $\varepsilon>0$，$c>0$，有

$$P\left(\left|\int_a^b X_t \mathrm{d}B_t\right| > \varepsilon\right) \leqslant P\left(\int_a^b X_t^2 \mathrm{d}t > c\right) + \frac{c}{\varepsilon^2}.$$

(ⅷ) 设$\{X, X_n\} \subset \mathcal{L}$为满足当 $n\to\infty$时

$$\int_a^b |X_n(t) - X(t)|^2 \mathrm{d}t \xrightarrow{P} 0$$

的序列，那么，当 $n\to\infty$时，下列结论成立

$$\sup_{a\leqslant t\leqslant b} \left|\int_a^t X_n(s)\mathrm{d}B(s) - \int_a^t X(s)\mathrm{d}B(s)\right| \xrightarrow{P} 0.$$

(ⅸ) 对任意 $\lambda>0$，下列结论成立

$$P\left(\sup_{a\leqslant t\leqslant T} \left|\int_a^t X_s \mathrm{d}B_s\right| > \lambda\right) \leqslant \frac{1}{\lambda^2} E\int_a^b X_s^2 \mathrm{d}s.$$

证明 (ⅰ)～(ⅳ)的证明留给读者．下面给出被积函数为简单可预报过程时二次变差公式(ⅴ)的证明，不失一般性，设$[a, b]=[0, t]$，$I(t) = \int_0^t X_s \mathrm{d}B_s$．设 X_t 的表达式为

$$X_s = X_0 + \sum_{i=0}^{n-1} \widetilde{X}_i \mathbf{1}_{(t_i, t_{i+1}]}(s),$$

其中，$\widetilde{X}_i$ 为$\mathcal{F}_{t_i}$可测的随机变量，那么

222

$$[I,I]_t = \sum_{i=0}^{n-1} \left[\int_{t_i}^{\cdot} X_s \mathrm{d}B_s, \int_{t_i}^{\cdot} X_s \mathrm{d}B_s\right]_{t_i}^{t_{i+1}},$$

这里$[\cdot, \cdot]_a^b$ 表示$[a, b]$上的二次变差．对子区间$[t_i, t_{i+1}]$的任一分法 Π_i：$t_i = s_0^{(i)} < \cdots < s_{m_i}^{(i)} = t_{i+1}$，有

$$\sum_{j=0}^{m_i-1} [I(s_{j+1}^{(i)}) - I(s_j^{(i)})]^2 = \widetilde{X}_i \sum_{j=0}^{m_i-1} [B(s_{j+1}^{(i)}) - B(s_j^{(i)})]^2,$$

当$\|\Pi_i\| \to 0$ 时，上式 a. s. 收敛于$\widetilde{X}_i^2(t_{i+1} - t_i)$，因此得到

$$[I,I]_t = \sum_{i=0}^{n-1} \widetilde{X}_i(t_{i+1} - t_i) = \int_0^t X_s \mathrm{d}s.$$

其他公式的证明可参考 Friedman(1975).

下列两点是值得讨论的，因为它们非常重要． ■

鞅性质

由定理 6.3.12 的(ⅲ)知：$\int_0^t X_s \mathrm{d}B_s$ 是一个均值为 0，方差为 $\sigma_t^2 = \int_0^t EX_s^2 \mathrm{d}s$ 的鞅. 特别地，对 $s \leqslant t$，

$$E\left(\int_0^t X_r \mathrm{d}B_r \mid \mathcal{F}_s\right) = \int_0^s X_r \mathrm{d}B_r$$

成立. 这个性质对任意被积函数$\{X_n\}$均成立.

方差与二次变差

区分方差与二次变差是非常必要的. 由性质(ⅴ)知后者为

$$S_t^2(\omega) = \int_0^t X^2(s,\omega)\mathrm{d}s, \omega \in \Omega.$$

二次变差刻画了一条路径(对任一固定的 ω)上的风险度量，通常它是一个非常数的随机变量. 如果对某一路径，取一个很大的状态 X_s，则其风险很高. 方差 $\sigma_t^2 = \mathrm{Var}\left(\int_0^t X_s \mathrm{d}B_s\right) = E\left(\int_0^t X_s \mathrm{d}B_s\right)^2$ 满足

$$\sigma_t^2 = ES_t^2 = \int S(t,\omega)^2 \mathrm{d}P(\omega)$$

(由 Itô等距性)，即在概率测度 P 下，二次变差在所有路径上的平均就是方差.

最后，将介绍 Itô积分的推广作为本小节的结尾. 到目前为止，已经介绍的随机积分都是以布朗运动作为积分变量，被积过程的样本路径均为有界变差的. 现希望积分变量 X 是一个可以分解为 $X_t = X_0 + B_t + A_t$ 这种形式的过程，其中，B_t 是一个布朗运动，A_t 是一个有界变差过程. 这种过程属于**半鞅**类. 所谓半鞅就是能分解为 $X_t = X_0 + M_t + A_t$ 的过程，其中，M_t 是一个鞅，A_t 是一个有界变差过程. 对这种积分变量 X，只要积分 $\int H\mathrm{d}B$ 与 $\int H\mathrm{d}A$ 都有定义，自然想到定义 $\int H\mathrm{d}X$ 为

223

$$\int H\mathrm{d}X = \int H\mathrm{d}B + \int H\mathrm{d}A.$$

在使用经典 L_2 逼近方法引入随机 Itô积分时，需要可预报的被积函数 f 满足条件 $\int_0^T Ef^2(t)\,\mathrm{d}t < \infty$，当然这也不是绝对的，上一节就略述了一种更一般的逼近方法，它仅要求被积函数 f 是几乎必然有限且是 càdlàg(右连续且左极限存在)的.

记基本概率空间为$(\Omega, \mathcal{F}, P)$，用 L_0 表示几乎必然有限(即满足 $P(X<\infty)=1$)的随机变量 X 的集合. 还是以简单可预报函数

$$H_n(t) = H_0 \mathbf{1}_{\{0\}} + \sum_{i=0}^{n-1} H_i \mathbf{1}_{(T_i, T_{i+1}]}, t \in [0, T]$$

为出发点，现在$\mathcal{F}_i$可测的随机变量 H_i 是 L_0 空间中的元素，且区间$[0, T]$的分点 $0 = T_0 < T_1 < \cdots T_n = T$ 可能是随机的，但假定它们是停时，即 T_i 是$\mathcal{F}_i$可测的，显然确定性分法是它的一种特殊情况.

对任一适应的 càdlàag 过程 X，按传统的方法定义随机积分 $\int HdX$，即

$$I(H)=\int_0^T H_s dX_s=\sum_{i=0}^{n-1} H_i(X_{T_{i+1}}-X_{T_i}).$$

更一般地，对 $t\in[0, T]$，定义

$$I(H)_t=\int_0^t H_s dX_s=H_0X_0+\sum_{i=0}^{n-1} H_i(X_{t\wedge T_{i+1}}-X_{t\wedge T_i}),$$

记简单可预报函数的集合为$\mathcal{E}$. 因为对任一 càdlàag 过程 H，存在一个序列$\{H_n\}\subset\mathcal{E}$，使得 H_n 在任一区间$[0, t]$上依概率一致收敛于 H，且在$[0, t]\times\Omega$ 上一致收敛于 H，即

$$\|H_n-H\|_\infty=\sup_{s\in[0,t]}\sup_{\omega\in\Omega}|H_n(s,\omega)-H(s,\omega)|\to 0,\quad n\to\infty. \tag{6.6}$$

如果当 $n\to\infty$时，积分 $I(H_n)=\int H_n dX$ 依概率收敛于一个随机变量，记该随机变量为 $I(H)=\int HdX$，则称积分变量 X 为可行的. 如果 X 为半鞅，即 X 可分解为 $X=M+A$，其中 M 为鞅，A 为有界变差过程，则有下列重要结果.

224

定理 6.3.13 设 X 是一个半鞅，那么 I: $\mathcal{E}\to L_0$ 为连续的线性映射，其中，$\mathcal{E}$中的范数规定为(6.6)式中的 sup 范数，L_0 的度量可以为任一蕴涵依概率收敛的度量[⊖].

换句话说，对于半鞅 X，可以将随机积分中的被积函数 H 推广为简单可预报函数空间$\mathcal{E}$中的函数列$\{H_n\}$的极限，这是因为存在一个随机变量 $I^*\in L_0$(记其为$\int HdX$)使得对任意 $\varepsilon>0$，有

$$\lim_{n\to\infty}\|H_n-H\|_\infty\Rightarrow\lim_{n\to\infty}P\left(\left|\int H_n dX-\int HdX\right|>\varepsilon\right)=0.$$

由上面的讨论可以看出，区间$[0, T]$上的随机积分，其结果为一个随机变量. 对 $0\leqslant t\leqslant T$，要求得随机积分对应的过程，对简单可预报函数 H_n，可以定义

$$I(H_n)(t,\omega)=\sum_{i=0}^{n-1} H_i(\omega)(X_{\min(t,T_{i+1})}(\omega)-X_{\min(t,T_i)}(\omega)), t\in[0,T],\omega\in\Omega,$$

即在和式中仅取 $T_i\leqslant t$ 的被加数. 这样随机积分对应了一个映射 I，其定义域为空间$\mathcal{E}$，取值为路径是 Skorohod 空间 $D[0, T]$中元素的随机过程，其中 $D[0, T]$表示在$[0, T]$上右连续且左极限存在的函数全体，其上配备一致收敛范数.

6.4 二次协变差

从定理 6.1.6 可以看出布朗运动的样本二次变差依 L_2 收敛于终端时间，即

$$[B,B]_t=t.$$

结合例 6.3.6 得到的公式，则有

$$[B,B]_t=B_t^2-2\int_0^t B_s dBs.$$

⊖ Prohorov 度量蕴涵依概率收敛.

这就导出下列定义.

定义 6.4.1 设$\{X_t\}$，$\{Y_t\}$是两个过程，满足$\int X\mathrm{d}Y$与$\int Y\mathrm{d}X$均存在，则称

$$[X,Y]_t = X_tY_t - \int_0^t X\mathrm{d}Y - \int_0^t Y\mathrm{d}X, \quad x\in[0,T]$$

为**二次协变差过程**或**括号过程**.

225

下列分部积分公式是显然的.

推论 6.4.2(分部积分)

$$\int_0^t X\mathrm{d}Y = XY|_0^t - [X,Y]_t - \int_0^t Y\mathrm{d}X.$$

二次协变差可以通过求对应的样本二次协变差的L_2极限得到.

定义和命题 6.4.3 对任一满足$\|\Pi_n\|\to 0$的分法Π_n：$0=t_0<\cdots<t_n=t$，**样本二次协变差**

$$[X,Y]_n(t) = [X,Y]_{nt} = \sum_{i=0}^{n-1}(X_{t_{i+1}}-X_{t_i})(Y_{t_{i+1}}-Y_{t_i})$$

依L_2收敛于$[X,Y]_t$，且收敛对$t\in[0,T]$是一致的.

证明 由$(b-a)(d-c)=bd-ac-a(d-c)-(b-a)c$知，当$n\to\infty$时，有

$$\begin{aligned}[X,Y]_n(t) &= \sum_{i=0}^{n-1}X_{t_{i+1}}Y_{t_{i+1}} - \sum_{i=0}^{n-1}X_{t_i}Y_{t_i} \\ &\quad - \sum_{i=0}^{n-1}Y_{t_i}(X_{t_{i+1}}-X_{t_i}) - \sum_{i=0}^{n-1}X_{t_i}(Y_{t_{i+1}}-Y_{t_i}) \\ &\to X_tY_t - X_0Y_0 - \int_0^t Y\mathrm{d}X - \int_0^t X\mathrm{d}Y,\end{aligned}$$

这是因为前两个和式收敛于$X_{t_n}Y_{t_n}-X_{t_0}Y_{t_0}=X_tY_t-X_0Y_0$，第三项与第四项的和是关于逼近阶梯函数

$$X_n(t) = \mathbf{1}_{\{0\}}X_0 + \sum_{i=0}^{n-1}X_{t_i}\mathbf{1}_{(t_i,t_{i+1}]}(t),$$
$$Y_n(t) = \mathbf{1}_{\{0\}}Y_0 + \sum_{i=0}^{n-1}Y_{t_i}\mathbf{1}_{(t_i,t_{i+1}]}(t)$$

的 Itô积分. 由定理 6.3.12 知，收敛对$t\in[0,T]$是一致的. ■

6.5 Itô公式

设f是一个光滑函数，B_t是一个布朗运动，本节的目标是推导出随机过程$f(B_t)$的表达式.

226

对$[0,t]$的任一分法$0=t_0<t_1<\cdots<t_n=t$，显然有

$$f(B_t) - f(B_0) = \sum_{i=0}^{n-1}[f(B_{t_{i+1}}) - f(B_{t_i})].$$

对和式的第i项使用 Taylor 公式，只取线性项及二次项，则有

$$f(B_{t_{i+1}})-f(B_{t_i})=f'(B_{t_i})(B_{t_{i+1}}-B_{t_i})+\frac{1}{2}f''(B_{t_i})(B_{t_{i+1}}-B_{t_i})^2+R.$$

略去余项 R，两边求和得到

$$f(B_t)-f(B_0)\approx\sum_{i=0}^{n-1}f'(B_{t_i})(B_{t_{i+1}}-B_{t_i})+\frac{1}{2}\sum_{i=0}^{n-1}f''(B_{t_i})(B_{t_{i+1}}-B_{t_i})^2.$$

上式的第一项(依 L_2)收敛于 Itô积分 $\int_0^t f'(B_s)\mathrm{d}B_s$，由定理 6.1.7 知上式的第二项收敛于 Taylor 展开式的二次项 $\int_0^t f''(B_s)\mathrm{d}s$. 下面将给出余项的一个猜想，并给出严谨的证明. 为此，先回顾一下下列形式的 Taylor 公式.

设 $f\in C^2$(二次连续可微函数空间)，则存在介于 x 与 y 之间的一点 ξ，使得

$$f(y)=f(x)+f'(x)(y-x)+\frac{1}{2}f''(\xi)(y-x)^2$$

成立. 上式可以改写为

$$f(y)=f(x)+f'(x)(y-x)+\frac{1}{2}f''(x)(y-x)^2+R_x,$$

其中余项为

$$R_x=\frac{1}{2}(f''(\xi)-f''(x))(y-x)^2.$$

推论 6.5.1(Itô公式)　设 $f:\mathbf{R}\to\mathbf{R}$ 是一个二阶连续可微函数，则对任意 $t>0$，有

$$f(B_t)-f(B_0)=\int_0^t f'(B_s)\mathrm{d}B_s+\frac{1}{2}\int_0^t f''(B_s)\mathrm{d}s.$$

证明　对$[0, t]$的任一满足$\|\Pi_n\|\to 0\quad(n\to\infty)$的分法 Π_n：$0=t_0<\cdots<t_n=t$，对表达式

227

$$f(B_t)-f(B_0)=\sum_{i=0}^{n-1}f(B_{t_{i+1}})-f(B_{t_i})$$

右边的每一项使用 f 的 Taylor 展开式，则有

$$f(B_{t_{i+1}})-f(B_{t_i})=f'(B_{t_i})(B_{t_{i+1}}-B_{t_i})+\frac{1}{2}f''(B_{t_i})(B_{t_{i+1}}-B_{t_i})^2+R_{t_i},$$

其中，

$$R_{t_i}=\frac{1}{2}(f''(\xi_i)-f''(B_{t_i}))(B_{t_{i+1}}-B_{t_i})^2,$$

且 ξ_i 是介于 B_{t_i} 与 $B_{t_{i+1}}$ 之间的一个随机变量. 由 B_t 的连续性知，集合

$$K_\omega=\{B_t(\omega):t\in[0,T]\}$$

对每一个 $\omega\in\Omega$ 均为紧集，所以 f''在集合 K_ω 上是一致连续的. 设 $\varepsilon>0$，$\delta=\delta(\varepsilon)>0$ 满足

$$|B_s-B_r|<\delta\Rightarrow|f''(B_s)-f''(B_r)|<\varepsilon,\quad 0\leqslant r,s\leqslant t.$$

因为当 $n\to\infty$ 时，$\|\Pi_n\|\to 0$，所以存在 $n_0=n_0(\delta)$，使得当 $n\geqslant n_0$ 时有 $\|\Pi_n\|<\delta$，且 $\sup\limits_{n\geqslant n_0}\|\Pi_n\|\leqslant\delta$. 从而第 i 项的余项满足

$$|R_{t_i}| = \frac{1}{2}|f''(\xi_i) - f''(B_{t_i})|(B_{t_{i+1}} - B_{t_i})^2 \leqslant \frac{\varepsilon}{2}(B_{t_{i+1}} - B_{t_i})^2.$$

由定理 6.1.6 知在 L_2 及 a.s. 意义下

$$\sum_{i=0}^{n-1}|R_{t_i}| \leqslant \frac{\varepsilon}{2}\sum_{i=0}^{n-1}(B_{t_{i+1}} - B_{t_i})^2 \to \frac{\varepsilon}{2}t.$$ ■

▶**例 6.5.2** 考虑 $Y_t = \frac{B_t^2}{2}$，即 $f(x) = \frac{x^2}{2}$，$x \in \mathbf{R}$ 时的 $f(B_t)$.

由 Itô公式有

$$\begin{aligned} f(B_t) &= f(B_0) + \int_0^t f'(B_s)\mathrm{d}B_s + \frac{1}{2}\int_0^t f''(B_s)\mathrm{d}s \\ &= f(0) + \int_0^t B_s \mathrm{d}B_s + \frac{1}{2}\int_0^t 1\mathrm{d}s \\ &= \int_0^t B_s \mathrm{d}B_s + \frac{t}{2}. \end{aligned}$$

进一步有

$$\int_0^t B_s \mathrm{d}B_s = \frac{1}{2}(B_t^2 - t).$$ ◀

▶**例 6.5.3** 设 B_t 是一个关于滤子 $\mathcal{F}_t$ 的布朗运动，μ_t 与 σ_t 是 $\mathcal{F}_t$ 适应过程，且假定 $\sigma_t \geqslant 0$， 228

$$E\int_0^T \sigma_t^2 \mathrm{d}t < \infty.$$

则 Itô积分

$$X_t = \int_0^t \left(\mu_s - \frac{\sigma_s^2}{2}\right)\mathrm{d}s + \int_0^t \sigma_s \mathrm{d}B_s$$

存在. 且过程

$$S_t = S_0 \exp\left\{\int_0^t \left(\mu_s - \frac{\sigma_s^2}{2}\right)\mathrm{d}s + \int_0^t \sigma_s \mathrm{d}B_s\right\}, \quad t \in [0, T]$$

有定义. 注意到

$$S_t = f(X_t),$$

因为 $f(x) = S_0 \mathrm{e}^x = f'(x) = f''(x)$，由 Itô公式有

$$S_t = \int_0^t \mu_u S_u \mathrm{d}u + \int_0^t \sigma_u S_u \mathrm{d}B_u, \quad t \in [0, T].$$ ◀

6.6 Itô过程

下面介绍一类重要的随机过程.

定义 6.6.1 设 $B_t(t \geqslant 0)$ 是一个关于滤子 $\{\mathcal{F}_t\}$ 的布朗运动，$\{\mu_t\}$ 与 $\{\sigma_t\}$ 是 $\mathcal{F}_t$ 适应及 $D_T \otimes \mathcal{F}_T$ 可测过程，且几乎必然满足

$$\int_0^T |\mu_s|\mathrm{d}s < \infty, \qquad \int_0^T |\sigma_s|^2 \mathrm{d}s < \infty,$$

那么，Itô随机积分 $\int \sigma_s \mathrm{d}B_s$ 存在，称满足

$$X_t = X_0 + \int_0^t \mu_s \mathrm{d}s + \int_0^t \sigma_s \mathrm{d}B_s$$

的过程$\{X_t\}$为始于 X_0，**漂移**为 μ_t，**波动**为 σ_t 的 **Itô过程**.

通常将 Itô过程 X_t 的积分形式改写为下列微分形式

$$\mathrm{d}X_t = \mu_t \mathrm{d}t + \sigma_t \mathrm{d}B_t.$$

现在可以得出这样一个重要结论：几何布朗运动的对数是一个 Itô过程，且这个 Itô过程的漂移及波动均为常数.

229

▶**例 6.6.2**(几何布朗运动) 回忆一下，对固定的常数 $\mu \in \mathbf{R}$ 及 $\sigma > 0$，如果

$$S_t = S_0 \exp(\mu t + \sigma B_t),$$

则称$\{S_t\}$为一个几何布朗运动. 现设 $X_t = \ln S_t$，那么

$$X_t = X_0 + \mu t + \sigma B_t$$

是一个对所有 t，漂移 $\mu_t = \mu$，波动 $\sigma_t = \sigma$ 的 Itô过程，其微分形式为

$$\mathrm{d}X_t = \mu \mathrm{d}t + \sigma \mathrm{d}B_t.$$ ◀

▶**例 6.6.3** Itô过程

$$X_t = \int_0^t \mu_s \mathrm{d}s + \int_0^t \sigma_s \mathrm{d}B_s$$

的二次变差为

$$[X, X]_t = \int_0^t \sigma_s^2 \mathrm{d}s.$$

证明 为了简化记号，记

$$m_t = \int_0^t \mu_s \mathrm{d}s, \quad s_t = \int_0^t \sigma_s \mathrm{d}B_s.$$

设 Π_n：$0 = t_0 < \cdots < t_n = t$ 是$[0, T]$的一个分法，则有

$$\sum_{i=0}^{n-1} (X_{t_{i+1}} - X_{t_i})^2 = \sum_{i=0}^{n-1} (S_{t_{i+1}} - S_{t_i})^2 + \sum_{i=0}^{n-1} (m_{t_{i+1}} - m_{t_i})^2 + 2\sum_{i=0}^{n-1} (s_{t_{i+1}} - s_{t_i})(m_{t_{i+1}} - m_{t_i}).$$

由定理 6.3.12 知上式第一项 a. s. 收敛于 $\int_0^t \sigma_s^2 \mathrm{d}s$. 上式第二项不大于

$$\max_{0 \leqslant i \leqslant n-1} |m_{t_{i+1}} - m_{t_i}| \cdot \sum_{i=0}^{n-1} |m_{t_{i+1}} - m_{t_i}|$$

且

$$\sum_{i=0}^{n-1} |m_{t_{i+1}} - m_{t_i}| = \sum_{i=0}^{n-1} \left| \int_{t_i}^{t_{i+1}} \mu_s \mathrm{d}s \right| \leqslant \int_0^t |\mu_s| \mathrm{d}s, \tag{6.7}$$

230

由 $M: u \mapsto \int_0^u \mu_s \mathrm{d}s$ 的连续性及 M 在$[0, t]$上是一致连续的知：当$\|\Pi_n\| \to 0$ 时，

$$\max_{0 \leqslant i \leqslant n-1} |m_{t_{i+1}} - m_{t_i}| \to 0.$$

而最后一项不大于

$$2\max_{0\leqslant i\leqslant n-1}|s_{t_{i+1}}-s_{t_i}|.\sum_{i=0}^{n-1}|m_{t_{i+1}}-m_{t_i}|.$$

因为 s_t 是连续的，由公式(6.7)知本例结论成立. ◀

推论 6.6.4 设$\{X_t\}$是一个Itô过程，$F_t=[X, X]_t$ 是可微的，且 $F_t'=\sigma_t^2$，那么，对任一过程$\{Z_t\}$，只要它关于$[X, X]_t$ 的积分存在，则其积分为

$$\int_0^t Z_s\mathrm{d}[X,X]_s=\int_0^t Z_s\sigma_s^2\mathrm{d}s. \qquad \blacksquare$$

下面的定理介绍如何计算某些适应过程关于Itô过程的积分.

定理 6.6.5 设

$$X_t=X_0+\int_0^t\mu_s\mathrm{d}s+\int_0^t\sigma_s\mathrm{d}B_s$$

是一个Itô过程，$\{Z_t\}$是一个适应过程，如果对任意 $t\in[0, T]$，均有

$$E\int_0^t Z_s^2\sigma_s^2\mathrm{d}s<\infty,\quad \int_0^t|Z_s\mu_s|\mathrm{d}s<\infty$$

成立，那么

$$\int_0^t Z_s\mathrm{d}X_s=\int_0^t Z_s\mu_s\mathrm{d}s+\int_0^t Z_s\sigma_s\mathrm{d}B_s,\quad 0\leqslant t\leqslant T.$$

证明 在$\{Z_t\}$，$\{\mu_t\}$，$\{\sigma_t\}$关于分法 $0=t_0<\cdots<t_n=T$ 均为简单可预报过程的情形下，给出该定理的证明. 因为当 $t\in(t_i, t_{i+1}]$时 $\mu_t=\mu_{t_i}$，$\sigma_t=\sigma_{t_i}$，所以

$$\begin{aligned}X_{t_{i+1}}-X_{t_i}&=\int_{t_i}^{t_{i+1}}\mu_s\mathrm{d}s+\int_{t_i}^{t_{i+1}}\sigma_s\mathrm{d}B_s\\&=\mu_{t_i}(t_{i+1}-t_i)+\sigma_{t_i}(B_{t_{i+1}}-B_{t_i}),\end{aligned}$$

231

从而得到

$$\begin{aligned}\int_0^t Z_s\mathrm{d}X_s&=\sum_{i=0}^{n-1}Z_{t_i}(X_{t_{i+1}}-X_{t_i})\\&=\sum_{i=0}^{n-1}Z_{t_i}\mu_{t_i}(t_{i+1}-t_i)+\sum_{i=0}^{n-1}Z_{t_i}\sigma_{t_i}(B_{t_{i+1}}-B_{t_i})\\&=\int_0^t Z_s\mu_s\mathrm{d}s+\int_0^t Z_s\sigma_s\mathrm{d}B_s.\end{aligned} \qquad \blacksquare$$

下面建立Itô过程的Itô公式.

定理 6.6.6(Itô过程的Itô公式) 设$\{X_t\}$是一个Itô过程，$f(t, x)$有连续偏导数 $f_t(t, x)$, $f_x(t, x)$，$f_{xx}(t, x)$，那么对 $t\in[0, T]$，有

$$\begin{aligned}f(t,X_t)&=f(0,X_0)+\int_0^t f_t(s,X_s)\mathrm{d}s+\int_0^t f_x(s,X_s)\mathrm{d}X_s+\frac{1}{2}\int_0^t f_{xx}(s,X_s)\mathrm{d}[X,X]_s\\&=f(0,X_0)+\int_0^t\left[f_t(s,X_s)+f_x(s,X_s)\mu_s+\frac{1}{2}f_{xx}(s,X_s)\sigma_s^2\right]\mathrm{d}s+\int_0^t f_x(s,X_s)\sigma_s\mathrm{d}B_s,\end{aligned}$$

换句话说，$f(t, X_t)$仍是Itô过程，其漂移为

$$\widetilde{\mu}_s = f_t(s,X_s) + f_x(s,X_s)\mu_s + \frac{1}{2}f_{xx}(s,X_s)\sigma_s^2,$$

波动为

$$\widetilde{\sigma}_s = f_x(s,X_s)\sigma_s.$$

定理 6.6.6 说明：一个 Itô过程的任意光滑函数仍是 Itô过程，且由 X_t 映射得到的新过程可能依赖于时间坐标 t.

▶**例 6.6.7** 考虑几何布朗运动

$$S_t = S_0\exp(\mu t + \sigma B_t),$$

其中，$\mu\in\mathbf{R}$，$\sigma>0$ 为常数．例 6.6.2 已经证明 $X_t=\log S_t$ 是一个漂移 $\mu_s=\mu$，波动 $\sigma_s=\sigma$ 的 Itô过程．现对 $S_t=f(X_t)$ 使用 Itô公式，其中 $f(x)=\mathrm{e}^x$，则 $S_t=f(X_t)$ 也是一个 Itô过程，其漂移为

232

$$\begin{aligned}\widetilde{\mu}_s &= f_x(X_s)\mu_s + \frac{1}{2}f_{xx}(X_s)\sigma_s^2\\ &= S_s\mu_s + \frac{1}{2}S_s\sigma_s^2\\ &= \left(\mu + \frac{\sigma^2}{2}\right)S_s\end{aligned}$$

波动为 $\widetilde{\sigma_s}=f_x(X_s)\sigma_s=S_s\sigma$，且

$$S_t = S_0 + \left(\mu + \frac{\sigma^2}{2}\right)\int_0^t S_r\mathrm{d}r + \sigma\int_0^t S_r\mathrm{d}B_r.$$

其微分形式为

$$\mathrm{d}S_t = \left(\mu + \frac{\sigma^2}{2}\right)S_t\mathrm{d}t + \sigma S_t\mathrm{d}B_t.$$

我们已经知道，几何布朗运动有表达式

$$S_t = S_0 + \left(\mu + \frac{\sigma^2}{2}\right)\int_0^t S_r\mathrm{d}r + \sigma\int_0^t S_r\mathrm{d}B_r.$$

在上式中，如果 $\mu=-\frac{\sigma^2}{2}$，则漂移项为 0，Itô过程变为 $S_t = S_0 + \sigma\int_0^t S_r\mathrm{d}B_r$，它是一个鞅.

推论 6.6.8 在几何布朗运动

$$S_t = S_0\exp(\mu t + \sigma B_t)$$

中，如果 $\mu=-\frac{\sigma^2}{2}$，那么该几何布朗运动是一个鞅. ◀

▶**例 6.6.9**(广义布朗运动) 设 μ_t，σ_t 是两个适应过程，且

$$\int_0^T|\mu_t|\mathrm{d}t<\infty,\qquad \int_0^T\sigma_t^2\mathrm{d}t<\infty,$$

几乎必然成立．过程 S_t 定义如下

233

$$S_t = S_0\exp\left\{\int_0^t\left(\mu_r - \frac{\sigma_r^2}{2}\right)\mathrm{d}r + \int_0^t\sigma_r\mathrm{d}B_r\right\},\quad t\in[0,T].$$

那么，

$$X_t = \log S_t = \log S_0 + \int_0^t \left(\mu_r - \frac{\sigma_r^2}{2}\right) \mathrm{d}r + \int_0^t \sigma_r \mathrm{d}B_r, \quad t \in [0,T]$$

是一个 Itô过程，其漂移 $\overline{\mu}_t$ 和波动 $\overline{\sigma}_t$ 分别为

$$\overline{\mu}_t = \mu t - \frac{\sigma_t^2}{2}, \quad \overline{\sigma}_t = \sigma_t.$$

因为 $S_t = \mathrm{e}^{X_t}$，所以由 Itô公式知 S_t 也是一个 Itô过程，其漂移

$$\begin{aligned}\widetilde{\mu}_t &= f_x(X_t)\overline{\mu}_t + \frac{1}{2} f_{xx}(X_t)\sigma_t^2 \\ &= \mathrm{e}^{X_t}\left(\mu_t - \frac{\sigma_t^2}{2}\right) + \frac{\sigma_t^2}{2} S_t \\ &= S_t \mu_t,\end{aligned}$$

波动

$$\widetilde{\sigma}_t = f_x(X_t)\sigma_t = S_t\sigma_t.$$

从而得到下列表达式

$$S_t = S_0 + \int_0^t \mu_r S_r \mathrm{d}r + \int_0^t \sigma_r S_r \mathrm{d}B_r, \quad t \in [0,T].$$

换句话说，$\{S_t\}$是随机微分方程

$$\mathrm{d}S_t = \mu_t S_t \mathrm{d}t + \sigma_t S_t \mathrm{d}B_t$$

的解. ◀

▶**例 6.6.10**(Vasicek 利率模型) 过程

$$X_t = \mathrm{e}^{-bt} X_0 + \frac{a}{b}(1 - \mathrm{e}^{-bt}) + \sigma \mathrm{e}^{-bt} \int_0^t \mathrm{e}^{bs} \mathrm{d}B_s, \quad t \in [0,T] \tag{6.8}$$

常用来表示利率模型，其中 X_0 是一个固定的初始值，$a \geqslant 0$ 和 b，$\sigma > 0$ 是参数. 这个过程经中间过程

$$Z_t = \int_0^t \mathrm{e}^{bs} \mathrm{d}B_s, \quad t \in [0,T]$$

与布朗运动相关.

记 f：$[0, T] \times \mathbf{R} \to \mathbf{R}$ 为

$$f(t,x) = \mathrm{e}^{-bt} X_0 + \frac{a}{b}(1 - \mathrm{e}^{-bt}) + \sigma \mathrm{e}^{-bt} x.$$

234

其偏导数为

$$\begin{aligned}f_t(s,x) &= -b\mathrm{e}^{-bs} X_0 + a\mathrm{e}^{-bs} - \sigma b \mathrm{e}^{-bs} x, \\ f_x(s,x) &= \sigma \mathrm{e}^{-bs}, \\ f_{xx}(s,x) &= 0.\end{aligned}$$

那么 $X_t = f(t, Z_t)$是一个 Itô过程，其漂移和波动分别为

$$\widetilde{\mu}_s = f_t(s,Z_s) + \frac{1}{2} f_{xx}(s,Z_s)\mathrm{e}^{2bs} = a - bf(s,Z_s) = a - bX_s$$

$$\widetilde{\sigma}_s = f_x(s,Z_s)\mathrm{e}^{bs} = \sigma.$$

因此，

$$X_t = X_0 + \int_0^t (a - bX_s)\mathrm{d}s + \sigma B_t,$$

即

$$\mathrm{d}X_t = (a - bX_t)\mathrm{d}t + \sigma \mathrm{d}B_t.$$

因为 X_t 出现在等式两边，所以这是一个随机微分方程. (6.8)式是这个随机微分方程的显式解. (6.8)式对模块有直观的解释，其漂移项依赖于现值. 对照几何布运动，在几何布朗运动中漂移为常数，而这个模型的漂移为 $a-bX_t$，$b>0$，如果 X_t 的值很大，则有一个局部减小的趋势. 如果 $X_t<a/b$，则漂移为正，如果 $X_t>a/b$，则漂移为负. ◀

当 $a=0$ 时，对应的模型为 **Langevin 随机微分方程**

$$\mathrm{d}X_t = -bX_t\mathrm{d}t + \sigma \mathrm{d}B_t,$$

其解为 **Ornstein-Unlenbeck 过程**

$$X_t = \mathrm{e}^{-bt}X_0 + \sigma\mathrm{e}^{-bt}\int_0^t \mathrm{e}^{bs}\mathrm{d}B_s, \quad t \geqslant 0.$$

当 $b>0$ 时，它是一个均值为 $E(X_t)=\mathrm{e}^{-bt}X_0 \to 0(t\to 0)$ 的高斯过程. 下面计算自协方差函数

$$\gamma_X(s,t) = \mathrm{Cov}(X_s, X_t), \quad s \leqslant t,$$

不妨假定 $X_0=0$. 分别用和式

$$X_s^{(n)} = \sigma\mathrm{e}^{-bs}\sum_{i=1}^{n}\mathrm{e}^{bs_{i-1}}[B_{s_i} - B_{s_{i-1}}]$$

与和式

235

$$X_t^{(n)} = \sigma\mathrm{e}^{-bt}\sum_{i=1}^{n+m}\mathrm{e}^{bs_{i-1}}[B_{s_i} - B_{s_{i-1}}]$$

去逼近 X_s 与 X_t，其中 $s_0<\cdots<s_{n+m}$ 是$[0,\ t]$的一个分法，$s_0<\cdots<s_n$ 是$[0,\ s]$的一个分法. 显然当 $n\to\infty$时，两个和式分别依 L_2 收敛于 X_s 与 X_t. 由内积的连续性知

$$\gamma_X(s,t) = \lim_{n\to\infty}\mathrm{Cov}(X_s^{(n)}, X_t^{(n)}),$$

且

$$\begin{aligned}\mathrm{Cov}(X_s^{(n)}, X_t^{(n)}) &= \sigma^2\mathrm{e}^{-b(s+t)}\sum_{i=1}^{n}\mathrm{e}^{2bs_{i-1}}(s_i - s_{i-1}) \\ &\to_{n\to\infty} \sigma^2\mathrm{e}^{-b(s+t)}\int_0^t \mathrm{e}^{2bs}\mathrm{d}s \\ &= \frac{\sigma^2}{2b}(\mathrm{e}^{-b(s-t)} - \mathrm{e}^{-b(s+t)}).\end{aligned}$$

特别地，当 $t\to\infty$时，

$$\mathrm{Var}(X_t) = \frac{\sigma^2}{2b}(1-\mathrm{e}^{-2bt}) \to \frac{\sigma^2}{2b}.$$

因为$\{X_t\}$是高斯过程，因此它是一个平稳且严平稳的过程，对独立于$\{B_t: t\geqslant 0\}$的随机初始条件

$$X_0 \sim N\left(0, \frac{\sigma^2}{2b}\right), \tag{6.9}$$

由 $\mathrm{Cov}(\mathrm{e}^{-bs}X_0, \mathrm{e}^{-bt}X_0)=\frac{\sigma^2}{2b}\mathrm{e}^{-b(s+t)}$ 得到

$$\mathrm{Cov}(X_s, X_t) = \frac{\sigma^2}{2b}\mathrm{e}^{-b|t-s|}, \quad s,t \geqslant 0, \tag{6.10}$$

它是 $|t-s|$ 的函数，这也印证了 $\{X_t\}$ 是平稳过程. ◀

命题 6.6.11 Langevin 随机微分方程有一个由初始值(6.9)给出的严平稳解. 这个解有均值 0 和自协方差函数(6.10).

6.7 扩散过程及遍历性

称随机微分方程

$$\mathrm{d}X_t = \mu(X_t)\mathrm{d}t + \sigma(X_t)\mathrm{d}B_t, X_0 = x_0, t \geqslant 0$$

的解 $\{X_t\}$ 为一个 **Itô扩散**. 前面使用 Itô公式已经验证了该方程的解为布朗运动、系数 $\mu(x)$ 及 $\sigma(x)$ 的函数. 通常情况下，称具有这种性质的解为方程的**强解**. 如果系数 $\mu(x)$ 及 $\sigma(x)$ 满足全局 Lipschitz 条件，即存在常数 L，使得

$$|\mu(x) - \mu(y)| + |\sigma(x) - \sigma(y)| \leqslant L|x - y|$$

236

成立，则方程必存在强解. 这种解可能是平稳的，Ornstein-Unlenbeck 就给出了一个具体的例子.

称一个扩散为**遍历的**，如果存在一个测度 π，称这个测度为**不变(概率)测度**，使得对所有满足 $\int |h(x)|\mathrm{d}\pi(x) < \infty$ 的函数 h，均有

$$\frac{1}{t}\int_0^t h(X_s)\mathrm{d}s = \int h(x)\mathrm{d}\pi(x)$$

几乎必然成立.

任一强解必是遍历的. 事实上，定义标度密度函数为

$$s(y) = \exp\left(-2\int_0^y \frac{\mu(x)}{\sigma^2(x)}\mathrm{d}x\right), \quad -\infty < y < \infty,$$

并假定标度函数

$$S(x) = \int_0^x s(y)\mathrm{d}y,$$

满足：当 $x \to \pm\infty$ 时，$S(x) \to \pm\infty$.

引入规范化常数

$$K = \int \frac{1}{\sigma^2(x)s(x)}\mathrm{d}x < \infty.$$

则唯一的平稳测度由**不变密度**

$$\mathrm{d}\mu(x) = \frac{1}{K\sigma^2(x)}\exp\left(2\int_0^x \frac{\mu(t)}{\sigma^2(t)}\mathrm{d}t\right)\mathrm{d}x$$

给出．进一步，如取初始条件

$$X_0 \sim \mu,$$

则对应的强解是严平稳的，通常记这个严平稳且遍历的解为$\{X_t\}$.

▶**例 6.7.1** 对 Ornstein-Unlenbeck 过程，已经知道其唯一的平稳测度是均值为 0，方差为$\frac{\sigma^2}{2b}$的正态分布，即 μ 满足

$$\mu((a,b]) = \int_a^b \varphi_{\left(0,\frac{\sigma^2}{2b}\right)}(x)\mathrm{d}x, \quad a \leqslant b.$$ ◀

几何布朗运动、Ornstein-Unlenbeck 过程及 CRI 模型均是参数为 α，β，σ 及 γ 的随机微分方程族

$$\mathrm{d}X_t = (\alpha + \beta X_t)\mathrm{d}t + \sigma X_t^\gamma \mathrm{d}B_t$$

中某个特殊方程的解．在金融中称这个随机微分方程族为 Chan-Karolyi-Longstaff-Sanders

237 (CKLS)模型.

下面是该模型的一些特殊情形及相应的名称.

(i) Merton 模型：$\mathrm{d}X_t = \alpha\mathrm{d}t + \sigma\mathrm{d}B_t$.

(ii) Vasicek 模型：$\mathrm{d}X_t = (\alpha + \beta X_t)\mathrm{d}t + \sigma\mathrm{d}B_t$.

(iii) CIR 平方根模型：$\mathrm{d}X_t = (\alpha + \beta X_t)\mathrm{d}t + \sigma\sqrt{X_t}\mathrm{d}B_t$.

(iv) Brennon-Schwartz 模型：$\mathrm{d}X_t = (\alpha + \beta X_t)\mathrm{d}t + \sigma X_t \mathrm{d}B_t$.

(v) CIR 变利率模型：$\mathrm{d}X_t = \sigma X_t^{3/2}\mathrm{d}B_t$.

(vi) CEV(不变方差弹性)模型：$\mathrm{d}X_t = \beta X_t \mathrm{d}t + \sigma X_t^\gamma \mathrm{d}B_t$.

6.8 数值逼近与统计估计

考虑一个非时齐的 Itô扩散

$$X_t = \int_0^t \mu(s, X_s)\mathrm{d}s + \int_0^t \sigma(s, X_s)\mathrm{d}B_s, \quad t \geqslant 0,$$

其中，$\mu(t, x)$，$\sigma(t, x)$是确定性函数，B_t 是布朗运动．微分形式为

$$\mathrm{d}X_t = \mu(t, X_t)\mathrm{d}t + \sigma(t, X_t)\mathrm{d}B_t. \tag{6.11}$$

一般情况下，$\{X_t\}$的显式解求不出来，所以只能求它的数值逼近，即求出一个离散时间$t\in\{t_1, t_2, \cdots\}$的过程，使之在某种意义下近似于$\{X_t\}$．注意：现实中是不能观测一个过程的所有路径的，一个连续时间的模型是理想化的数学描述，能观测到的仅是离散时间点 t_1，t_2，…的瞬象．假设这些离散时间点的步长相等，都为 Δ_n，则

$$t_i = t_{ni} = t_0 + i\Delta_n, \quad i = 1, 2, \cdots,$$

记过程$\{X_t, t\geqslant 0\}$在这些点的样本(瞬象)为

$$X_i^{(n)} = X_{t_i}, \quad i = 1, \cdots, T,$$

则当 Δ_n 充分小时有下列近似

$$\int_{t_{i-1}}^{t_i} \mu(s, X_s) \mathrm{d}s \approx \mu(t_{i-1}, X_{t_{i-1}})(t_i - t_{i-1})$$

及

$$\int_{t_{i-1}}^{t_i} \sigma(s, X_s) \mathrm{d}B_s \approx \sigma(t_{i-1}, X_{t_{i-1}})(B_{t_{i-1}} - B_{t_i}),$$

其中，对所有 i，$t_i - t_{i-1} = \Delta_n$，$B_{t_{i-1}} - B_{t_i} \sim \sqrt{\Delta} N(0, 1)$.

238

建立在这种近似基础上，就得到 SDE(6.1)的 **Euler 逼近方法**

$$\Delta X_i^{(n)} = \mu(t_0 + (i-1)\Delta_n, X_{i-1}^{(n)})\Delta_n + \sigma(t_0 + (i-1)\Delta_n, X_{i-1}^{(n)}) \sqrt{\Delta_n} \varepsilon_i, \tag{6.12}$$

其中，

$$\Delta X_i^{(n)} = X_i^{(n)} - X_{i-1}^{(n)}$$

且 ε_i 是 i. i. d. 的标准正态随机变量.

对时齐扩散

$$\mathrm{d}X_t = \mu(X_t)\mathrm{d}t + \sigma(X_t)\mathrm{d}B_t,$$

对应于(6.12)的 Euler 逼近方法为

$$\Delta X_i^{(n)} = \mu(X_{i-1}^{(n)})\Delta_n + \sigma(X_{i-1}^{(n)}) \sqrt{\Delta_n} \varepsilon_i, \tag{6.13}$$

其中，$X_{t_i}^{(n)} = X_i^{(n)}$ 是时刻 $t_i = t_0 + i\Delta_n$ 点的近似. 在网格点之间，该方法使用分段常数插值或线性近似来构造 $X_i^{(n)}$.

不难看出，模型(6.13)属于一般非参数回归模型(3.1)的类型，在第 9 章中，将进一步讨论这种模型.

下面的结果表明：在系数一个满足 Lipschitz 条件，一个满足线性增长的条件下，上述近似有强收敛率 $\Delta_n^{1/2}$.

定理 6.8.1 设 $E|X_0|^2 < \infty$，$\|X_0 - X_0^{(n)}\|_2 = O(\delta^{1/2})$. 如果对任意 s，$t \in [0, T]$及 x，$y \in \mathbf{R}^d$，均有 Lipschitz 条件

$$|\mu(t,x) - \mu(t,y)| + |\sigma(t,x) - \sigma(t,y)| = K_1 |x - y|,$$

线性增长条件

$$|a(t,x)| + |b(t,x)| \leqslant K_2(1 + |x|),$$

及

$$|a(s,x) - a(t,x)| + |b(s,x) - b(t,x)| \leqslant K_3(1 + |x|)|s - t|^{1/2}$$

成立，其中 K_1，K_2，K_3 不依赖于 Δ_n，那么，Euler 逼近 $X_t^{(n)}$ 满足下列一致估计

$$\sup_{t \in [0,T]} E(|X_t - X_t^{(n)}|) = O(\Delta_n^{1/2}).$$

6.9 评注与延伸阅读

本章的内容在许多专著及教材上都可找到. 关于有界变差函数的非随机积分可阅读 Lang(1993)，随机积分的经典著作是 Friedman(1975). 关于 Itô积分 L_2 理论的简洁介绍可参考 Øksendal(2003). 比本章内容更详细的介绍可阅读 Klebaner(2005)，本章 Itô公式的

239

证明选自 Klenke(2008). Shreve(2004)直观且生动的处理方式启发了作者写作本章内容. 对那些不喜欢对每一个定理都给出严格数学证明的读者，我们推荐综合介绍这部分内容的书 Grigoriu(2002)及简洁版教材 Mikosch(1998)作为辅助读物. Protter(2005)是一本随机分析的高等教材，本章中许多简洁的证明，比如分部积分公式的证明，都是选自该书，关于如何将随机积分的被积过程推广到右连左极过程的内容也是取自该书，读者也可参考 Kurtz 和 Protter(1996)、Jacod 和 Shiryaev(2003). CKLS 模型选自 Chen 等(1992). 定理 6.8.1 选自 Kloeden 和 Platen(1992)中定理 10.2.2.

参考文献

Chan K.C., Longstaff F.A. and Sanders A.B. (1992) An empirical comparison of alternative models of the short-term interest rate. *Journal of Finance* **47**(3), 1209–1227.

Friedman A. (1975) *Stochastic Differential Equations and Applications. Vol. 1.* Academic Press [Harcourt Brace Jovanovich Publishers], New York. Probability and Mathematical Statistics, Vol. 28.

Grigoriu M. (2002) *Stochastic Calculus: Applications in Science and Engineering*. Birkhäuser Boston Inc., Boston, MA.

Jacod J. and Shiryaev A.N. (2003) *Limit Theorems for Stochastic Processes*. vol. 288 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]* 2nd edn. Springer-Verlag, Berlin.

Klebaner F.C. (2005) *Introduction to Stochastic Calculus with Applications*. 2nd edn. Imperial College Press, London.

Klenke A. (2008) *Probability Theory*. Universitext. Springer-Verlag London Ltd., London. A comprehensive course, Translated from the 2006 German original.

Kloeden P.E. and Platen E. (1992) *Numerical Solution of Stochastic Differential Equations*. vol. 23 of *Applications of Mathematics (New York)*. Springer-Verlag, Berlin.

Kurtz T.G. and Protter P.E. (1996) *Weak Convergence of Stochastic Integrals and Differential Equations*. vol. 1627 of *Lecture Notes in Math.*. Springer, Berlin.

Lang S. (1993) *Real and Functional Analysis*. vol. 142 of *Graduate Texts in Mathematics*. 3rd edn. Springer-Verlag, New York.

Mikosch T. (1998) *Elementary Stochastic Calculus—with Finance in View*. vol. 6 of *Advanced Series on Statistical Science & Applied Probability*. World Scientific Publishing Co. Inc., River Edge, NJ.

Øksendal B. (2003) *Stochastic Differential Equations: An Introduction with Applications*. Universitext 6th edn. Springer-Verlag, Berlin.

Protter P.E. (2005) *Stochastic Integration and Differential Equations*. vol. 21 of *Stochastic Modelling and Applied Probability*. Springer-Verlag, Berlin. 2nd edition. Version 2.1, Corrected third printing.

Shreve S.E. (2004) *Stochastic Calculus for Finance. II: Continuous-time Models*. Springer Finance. Springer-Verlag, New York.

第7章 Black-Scholes模型

Black-Scholes模型假定投资者或者把资金投资于股市或者储蓄在银行作为存款，而前者的价格是随机的. 对比迄今为止已经学习过的离散时间模型，两者最大的不同是本章的金融市场是用连续时间的随机过程建模并且允许连续交易. 因此，像股票价格或者交易策略等相关的量也都是连续时间的随机过程.

本章首先讨论银行账户支付已知且固定利率的经典模型公式，该模型是无套利、完备的. 并通过Girsanov定理所确定的唯一等价鞅测度推出了第1章已经讨论过的，关于欧式看涨期权无套利定价的著名Black-Scholes公式. 这个模型可以轻易地推广到股价过程的波动率仍是确定性的但依赖于时间的情形. 最后，本章简要地讨论了推广的Black-Scholes模型，在这类模型中股价波动率和利率都可能是随机过程.

7.1 模型和第一性质

为了简化在连续时间情况下的叙述，本章改变一些标准记号. 到现在为止，当使用双重指标时，第一个指标是时间指标，第二个指标是金融工具，也就是说S_{ti}表示t时刻第i个金融工具的价格. 从现在开始，改用记号S_{it}，因为以下经常用到一个固定工具关于时间的被积函数，在这种情况下，$\int_0^t S_{1u}\mathrm{d}u$比$\int_0^t S_{u1}\mathrm{d}u$更自然.

Black和Scholes推出的模型如下：给定一个固定利率r的银行账户$\{S_{0t}: t\in[0, T]\}$用于存款和借贷，同时给定一个风险资产，它是定义在带滤的概率空间$(\Omega, \mathcal{F}, \{\mathcal{F}_t\}, P)$上的可交易的股票$\{S_{1t}: t\in[0, T]\}$. 与前面各章一样，$P$表示实际概率测度. 假设过程$\{S_{0t}\}$和$\{S_{1t}\}$分别为

$$S_{0t}=\mathrm{e}^{rt},\quad t\in[0,T],$$
$$S_{1t}=S_{10}\exp\left(\left(\mu-\frac{\sigma^2}{2}\right)t+\sigma B_t\right),\quad t\in[0,T],$$

其中，$\mu\in\mathbf{R}$，$\sigma>0$和S_{10}为常数，$\{B_t:t\in[0, T]\}$是测度P下的标准布朗运动. 显然，$\{S_{1t}\}$是漂移参数为$\mu-\sigma^2/2$、波动为σ的几何布朗运动，且银行账户满足积分方程

$$S_{0t}=S_{00}+r\int_0^t S_{0u}\mathrm{d}u,\quad t\in[0,T],$$

或等价地

$$\mathrm{d}S_{0t}=rS_{0t}\mathrm{d}t,$$

其解的显式表达式为：

$$S_{0t}=\mathrm{e}^{rt},\quad t\in[0,T].$$

由例6.6.7知，$\{S_{1t}\}$有下列表达式：

$$S_{1t}=S_{10}+\mu\int_0^t S_{1u}\mathrm{d}u+\sigma\int_0^t S_{1u}\mathrm{d}B_u,\quad t\in[0,T],$$

即

$$dS_{1t}=\mu S_{1t}dt+\sigma S_{1t}dB_t.$$

对后一个模型方程可作如下解释：在一个很小的时间区间$[t, t+dt]$上，股票价格的变化近似于$\mu S_{1t}dt$，这是一个确定性的量，因为S_{1t}在t时刻是已知的．并且股价是被随机噪声项$\sigma S_{1t}\sqrt{dt}Z$扰动，其中

$$Z=\frac{B_{t+dt}-B_t}{\sqrt{dt}}\sim N(0,dt)$$

独立于S_{1t}．换句话说，相对的价格变化(即收益)被由一个正态分布的误差项扰动的线性趋势所刻画，即

$$\frac{dS_{1t}}{S_{1t}}\approx\mu dt+\sigma\sqrt{dt}Z.$$

通常习惯于用“年”作为时间单位，则μ是股价的年平均增长率，σ是年波动率．下面简要地讨论一下如何在实际中估计这些参数．注意到对数股票价格

242

$$\log S_{1t}=\log S_{10}+\left(\mu-\frac{\sigma^2}{2}\right)t+\sigma B_t$$

服从均值为$\log S_{10}+\left(\mu-\frac{\sigma^2}{2}\right)$，方差为$\sigma^2 t$的正态分布，且在时间长度 lag$\Delta$ 上的对数收益

$$R_t=R_{t\Delta}=\log S_{1,t+\Delta}-\log S_{1t}=\left(\mu-\frac{\sigma^2}{2}\right)\Delta+\sigma(B_{t+\Delta}-B_t)$$

是均值为$\left(\mu-\frac{\sigma^2}{2}\right)\Delta$，方差为$\Delta\sigma^2$的正态分布．从过程中以等距网格$t_i=i\Delta(i=1, \ldots, n)$抽样，相应的对数收益$R_1$，…，$R_n$是含有这些参数的独立同分布的正态分布．因此，波动参数σ有估计

$$\hat{\sigma}_n=\frac{sd_R}{\sqrt{\Delta}},$$

其中，

$$sd_R^2=\frac{1}{n-1}\sum_{t=1}^{n}(R_t-\bar{R}_n)^2,\quad \bar{R}_n=\frac{1}{n}\sum_{t=1}^{n}R_t,$$

$$\hat{\mu}_n=\frac{1}{\Delta}\bar{R}_n+\frac{sd_R^2}{2}\approx\frac{1}{\Delta}\bar{R}_n.$$

图 7-1 描述了 Black-Scholes 模型的模拟路径，其参数μ和σ由实际价格序列估计得出，实际价格序列也被加入图中作为参照.

现考虑贴现股价过程：

$$\begin{aligned}S_{1t}^*&=S_{0t}^{-1}S_{1t}\\&=e^{-rt}S_{1t}\\&=S_{10}\exp\left(\left(\mu-r-\frac{\sigma^2}{2}\right)t+\sigma B_t\right).\end{aligned}$$

那么$\{S_{1t}^*\}$是一个 Itô过程吗？注意到 S_{1t}^* 仍是一个几何布朗运动，所以马上得到下列的结论.

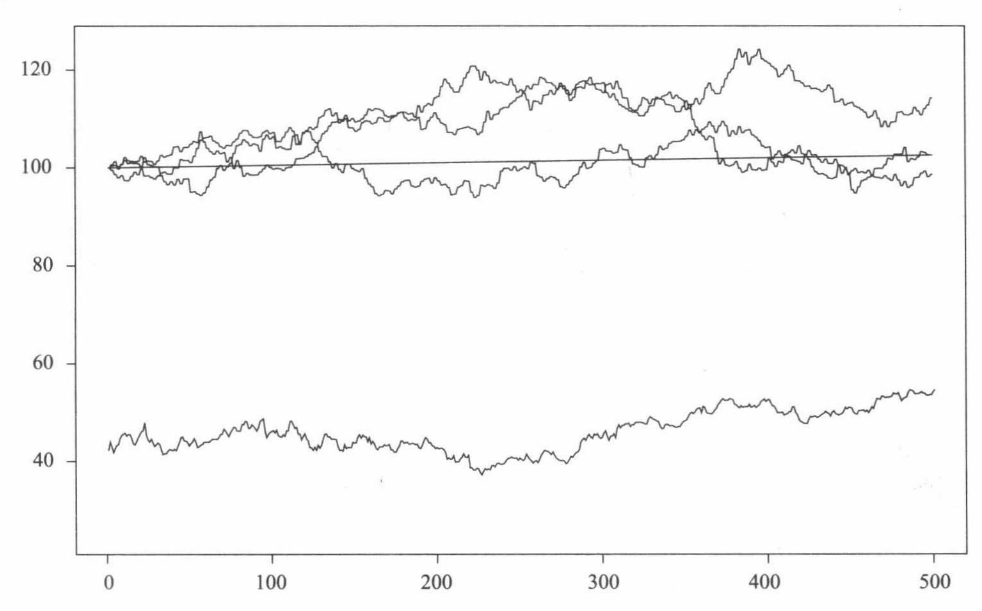


图 7-1 Black-Scholes 模型的模拟路径. 直线表示银行账户，始于 100 的曲线表示股票价格路径的三条独立复制. 参数由瑞信(Credit Suisse)公司的每日对数收益估计而来，其价格过程亦被加入作为参照，数据来源于统计软件 R 提供的免费欧洲股票市场数据集

引理 7.1.1 股价贴现过程$\{S_{1t}^*\}$是一个 Itô过程，

$$S_{1t}^* = S_{10} + (\mu - r)\int_0^t S_{1u}^* \mathrm{d}u + \sigma\int_0^t S_{1u}^* \mathrm{d}B_u,$$

或等价地

$$\mathrm{d}S_{1t}^* = (\mu - r)S_{1t}^* \mathrm{d}t + \sigma S_{1t}^* \mathrm{d}B_t.$$

在允许投资者能连续交易直到终端时刻 T 的条件下，需要在描述交易的连续时间过程中加入一些正则性条件. 首先，如同离散时间情形一样，要求在 t 时刻持有的股份数不能使用未来的信息. 其次，加入一个条件使得相关的(随机)积分是有定义的. 243

定义 7.1.2 **一个交易策略(资产投资组合)**是一个满足下列条件(ⅰ)～(ⅲ)的二维随机过程：

$$\varphi_t = (\alpha_t, \beta_t), \quad t \in [0, T],$$

其中，

(ⅰ) φ: $[0, T] \times \Omega \to \mathbf{R}^2$ 是$\mathcal{B}([0, T]) \otimes \mathcal{F}_t$ 可测的.

(ⅱ) φ 是$\mathcal{F}_t$ 适应的.

(ⅲ) $\int_0^T \alpha_t^2 \mathrm{d}t < \infty, \int_0^T |\beta_t| \mathrm{d}t < \infty.$

给定一个交易策略$\{(\alpha_t, \beta_t): t\in[0, T]\}$，$\alpha_t$ 表示风险资产在 t 时刻拥有的股份数，β_t 表示银行存款数. 条件(ⅲ)保证了随机积分

244

$$\int_0^t \beta_u \mathrm{d}S_{0u} = r\int_0^t \beta_u S_{0u}\mathrm{d}u, \quad \text{因为}(S_{0u}' = r\mathrm{e}^{ru})$$

及

$$\int_0^t \alpha_u \mathrm{d}S_{1u} = \mu\int_0^t \alpha_u S_{1u}\mathrm{d}u + \sigma\int_0^t \alpha_u S_{1u}\mathrm{d}B_u$$

有定义.

为了理解下面的定义，回顾一下数理金融中 Itô积分 $\int_0^t \alpha_u \mathrm{d}S_{1u}$ 的解释，理解这个积分的经济含义是非常重要的，它表示在可交易金融工具 S_{1u}(积分变量)里，一个随时间变化的持有量 α_u(被积函数)的价值过程. 在时刻 u 的持有量是 α_u，在无穷小时间段$[u, u+\Delta]$持有量变动的价值为 $\alpha_u \mathrm{d}S_{1u}(\approx\alpha_u(S_{1,u+\Delta}-S_{1u}))$. 对应于交易策略 $\varphi=(\alpha, \beta)$的价值过程为

$$V_t(\varphi) = \alpha_t S_{1t} + \beta_t S_{0t}, \quad t\in[0,T],$$

$V_0(\varphi)$是交易策略$\{\varphi_t\}$的初始成本.

定义 7.1.3 称一个交易策略$\{\varphi_t\}$为**自融资的**(self-financing)，如果其价值过程满足

$$V_t(\varphi) = V_0(\varphi) + \int_0^t \alpha_u \mathrm{d}S_{1u} + \int_0^t \beta_u \mathrm{d}S_{0u}, \quad t\in[0,T],$$

或等价地 $\mathrm{d}V_t(\varphi)=\alpha_t \mathrm{d}S_{1t}+\beta_t \mathrm{d}S_{0t}$.

一个自融资策略表明，资产组合中价值的变化是由于持有的股票和银行存款的价值变化分别引起的.

引理 7.1.4 假设 $\varphi=\{\varphi_t\}$是一个自融资交易策略，则其价值过程$\{V_t(\varphi)\}$是一个 Itô 过程，

$$V_t(\varphi) = V_0(\varphi) + \int_0^t \mu_V(u)\mathrm{d}u + \int_0^t \sigma_V(u)\mathrm{d}B_u, \quad t\in[0,T],$$

其漂移为

$$\mu_V(t) = \mu\alpha_t S_{1t} + r\beta_t S_{0t}$$

波动为

$$\sigma_V(t) = \sigma\alpha_t S_{1t}.$$

证明 因为 φ_t 是自融资的，所以 $V_t=V_t(\varphi)$满足

245

$$V_t = V_0 + \int_0^t \alpha_u \mathrm{d}S_{1u} + \int_0^t \beta_u \mathrm{d}S_{0u}, \tag{7.1}$$

其中

$$S_{0t} = \mathrm{e}^{rt} = S_{00} + \int_0^t rS_{0u}\mathrm{d}u$$

且

$$S_{1t} = S_{10} + \int_0^t \mu S_{1u}\mathrm{d}u + \int_0^t \sigma S_{1u}\mathrm{d}B_u$$

是 Itô过程. 下面计算(7.1)式中的随机积分. 注意到两种情况的积分变量都是 Itô积分，由

定理 6.6.5 知：

$$\int_0^t \alpha_u \mathrm{d}S_{1u} = \int_0^t \alpha_u \mu S_{1u} \mathrm{d}u + \int_0^t \alpha_u \sigma S_{1u} \mathrm{d}B_u$$

且

$$\int_0^t \beta_u \mathrm{d}S_{0u} = r\int_0^t \beta_u \mathrm{e}^{ru} \mathrm{d}u.$$

■

下列结论表明一个自融资策略的贴现价值过程具有特殊的结构.

推论 7.1.5 一个交易策略是自融资的，当且仅当

$$V_t^*(\varphi) = V_0(\varphi) + \int_0^t \alpha_u \mathrm{d}S_{1u}^*, \quad t \in [0,T],$$

其中，对 $t\in[0, T]$，

$$\int_0^t \alpha_u \mathrm{d}S_{1u}^* = \int_0^t (\mu - r)\alpha_u S_{1u}^* \mathrm{d}u + \int_0^t \sigma\alpha_u S_{1u}^* \mathrm{d}B_u.$$

证明 给出必要性的证明. 为了记号的简便，令 $V_t = V_t(\varphi)$，$V_t^* = V_t^*(\varphi) = \mathrm{e}^{-rt}V_t$. 由 e^{-rt} 是有界变差过程知括号过程 $[V, \mathrm{e}^{-r\cdot}]_t$ 为 0. 因此，

$$0 = [V, \mathrm{e}^{-r\cdot}]_t = V_u \mathrm{e}^{-ru}\big|_0^t - \int_0^t V_u \mathrm{d}(\mathrm{e}^{-ru}) - \int_0^t \mathrm{e}^{-ru} \mathrm{d}V_u.$$

246

其中，第一项为 $V_t^* - V_0^*$，从而

$$V_t^* = V_0^* + \int_0^t V_u \mathrm{d}(\mathrm{e}^{-ru}) + \int_0^t \mathrm{e}^{-ru} \mathrm{d}V_u.$$

现计算上式右边的积分项. 对每个 $t\in[0, T]$，有

$$\begin{aligned}\int_0^t \mathrm{e}^{-ru} \mathrm{d}V_u &= \int_0^t \mathrm{e}^{-ru}\mu_V(u)\mathrm{d}u + \int_0^t \mathrm{e}^{-ru}\sigma_V(u)\mathrm{d}B_u \\ &= \int_0^t (\mu\alpha_u S_{1u}^* + r\beta_u)\mathrm{d}u + \int_0^t \sigma\alpha_u S_{1u}^* \mathrm{d}B_u.\end{aligned}$$

进一步，由 $V_u = \alpha_u S_{1u} + \beta_u S_{0u}$ 且 $S_{0u} = \mathrm{e}^{ru}$，得

$$\begin{aligned}\int_0^t V_u \mathrm{d}(\mathrm{e}^{-ru}) &= -\int_0^t V_u r \mathrm{e}^{-ru} \mathrm{d}u \\ &= -\int_0^t (r\alpha_u S_{1u}^* + r\beta u)\mathrm{d}u.\end{aligned}$$

综上所述，可得：对任意 $t\in[0, T]$，有

$$V_t^* - V_0^* = \int_0^t (\mu - r)\alpha_u S_{1u}^* \mathrm{d}u + \int \sigma\alpha_u S_{1u}^* \mathrm{d}B_u,$$

这就完成了必要性的证明. ■

7.2 Girsanov 定理

Girsanov 定理提供这样一种方法：通过改变实际概率测度 P 使得贴现股价过程变成一个鞅. 在 Black-Scholes 模型中，对数股票价格是由带线性漂移项的布朗运动所驱动，因此

首要且关键的一步是解决这类过程的问题.

先介绍一些预备知识. 假设 $L\geqslant 0$ 是概率空间$(\Omega, \mathcal{F}, P)$中的随机变量且 $E(L)=1$，则可以用 L 作为 Radon-Nikodym 导数$\frac{\mathrm{d}Q}{\mathrm{d}P}$定义一个新的概率测度 Q，即

$$Q(A)=\int_A L\,\mathrm{d}P, \quad A\in\mathcal{F}. \tag{7.2}$$

任何由 P-密度定义的概率测度关于 P 都是绝对连续的. 另外，如果 L 是正的，即 $L(\omega)>0$ 对 $\forall\,\omega\in\Omega$，则 Q 等价于 P.

247

引理 7.2.1 假设 X:$(\Omega, \mathcal{F}, P)\to(\mathbf{R}, \mathcal{B})$服从 $N(\mu, \sigma^2)$分布，其中 $\mu\in\mathbf{R}$，$\sigma^2>0$. 定义

$$L(\omega)=\exp\left(-\frac{\mu}{\sigma^2}X(\omega)+\frac{\mu^2}{2\sigma^2}\right), \quad \omega\in\Omega.$$

则 $E(L)=1$，且在(7.2)式给定的概率测度 $\mathrm{d}Q=L\mathrm{d}P$ 下，X 服从 $N(0, \sigma^2)$分布.

证明 $L>0$ 是显然的. $E(L)=1$ 的证明留给读者. 下面计算 X 在概率测度 Q 下的特征函数，证明它就是 $N(0, \sigma^2)$分布的特征函数. 事实上，对任意的 t，如果在 P 下 $Y\sim N(0, \sigma^2)$，则有

$$\begin{aligned}
E_Q(\mathrm{e}^{\mathrm{i}tX}) &= E_P(L\mathrm{e}^{\mathrm{i}tX}) \\
&= E_P\exp\left(-\frac{\mu}{\sigma^2}X+\frac{\mu^2}{2\sigma^2}+\mathrm{i}tX\right) \\
&= \frac{1}{\sqrt{2\pi}\sigma}\int\exp\left(-\frac{2\mu}{2\sigma^2}x+\frac{\mu^2}{2\sigma^2}+\mathrm{i}tx-\frac{(x-\mu)^2}{2\sigma^2}\right)\mathrm{d}x \\
&= \frac{1}{\sqrt{2\pi}\sigma}\int\exp\left(-\frac{x^2}{2\sigma^2}+\mathrm{i}tx\right)\mathrm{d}x \\
&= E_P(\mathrm{e}^{\mathrm{i}tY}).
\end{aligned}$$

设$\{B_t: t\in[0, T]\}$如股票价格过程的定义中出现的那样是 P 下的标准布朗运动. 由推论 6.6.8 知，对任意固定的 $m\in R$，过程

$$L_t=\exp(-mB_t-\frac{m^2}{2}t), \quad t\in[0,T], \tag{7.3}$$

是一个正鞅且满足 $E(L_t)=1$. 因此，可以用 L_t 作为 P-密度通过(7.2)式去定义一个等价的概率测度. ■

引理 7.2.2 由 $\mathrm{d}Q=L_T\mathrm{d}P$ 确定的概率测度 Q，关于概率空间$(\Omega, \mathcal{F}_t, P)$的 P-密度为 L_t，其中$\{L_t: t\in[0, T]\}$由(7.3)式给出.

证明 只需证

$$Q(A)=\int_A L_t\mathrm{d}P, \quad 对所有\ A\in\mathcal{F}_t.$$

令 $A\in F_t$，则

$$\begin{aligned}
Q(A) &= E(\mathbf{1}_A L_T) \\
&= E(E(\mathbf{1}_A L_T\,|\,\mathcal{F}_t))
\end{aligned}$$

$$=E(\mathbf{1}_A E(L_T|\mathcal{F}_t))$$
$$=E(\mathbf{1}_A L_t),$$

得证. ■ 248

引理 7.2.3 设$\mathcal{A}\subset\mathcal{F}$是一个子$\sigma$域，$X$是一个随机变量，且满足

$$E(\mathrm{e}^{\mathrm{i}tX}|\mathcal{A})=\mathrm{e}^{-t^2\sigma^2/2},\quad t\in\mathbf{R}.$$

则X独立于$\mathcal{A}$且$X\sim N(0,\sigma^2)$.

证明 设$A\in\mathcal{A}$，注意到

$$\begin{aligned}E(\mathbf{1}_A\mathrm{e}^{\mathrm{i}tX})&=E(\mathbf{1}_A E(\mathrm{e}^{\mathrm{i}tX}|\mathcal{A}))\\&=E(\mathbf{1}_A\mathrm{e}^{-t^2\sigma^2/2})\\&=P(A)\mathrm{e}^{-t^2\sigma^2/2}\end{aligned}$$

对所有$t\in\mathbf{R}$均成立. 特别地，X的特征函数$\varphi_X(t)=\mathrm{e}^{-t^2\sigma^2/2}$(取$A=\Omega$)，所以$X\sim N(0,\sigma^2)$. 由假设知$X$在条件$A\in\mathcal{A}$下的条件分布也是$N(0,\sigma^2)$，即

$$P(X\leqslant x|A)=\Phi\left(\frac{x}{\sigma}\right),\quad x\in\mathbf{R}.$$

则对所有的$x\in\mathbf{R}$和$A\in\mathcal{F}$有

$$\begin{aligned}P(\{X\leqslant x\}\cap A)&=P(X\leqslant x|A).P(A)\\&=\Phi\left(\frac{x}{\sigma}\right)P(A).\end{aligned}$$

因为 Borel σ域$\mathcal{B}$由区间集$(-\infty,x]$，$x\in\mathbf{R}$生成，所以X和$\mathcal{A}$是独立的.

接下来将用到下列事实：随机变量$X\sim N(\mu,\sigma^2)(\mu\in\mathbf{R},\sigma^2>0)$的特征函数

$$\varphi_X(t)=E(\mathrm{e}^{\mathrm{i}tX})=\mathrm{e}^{\mathrm{i}t\mu-\sigma^2t^2/2},\quad t\in\mathbf{R},$$

在复数域内仍成立. 事实上，函数$\Phi_X(z)=\mathrm{e}^{\mathrm{i}z\mu-\sigma^2z^2/2}(z\in\mathbf{C})$是解析的，在实数域内与$\varphi_X$相等，且$\Phi_X$和$\varphi_X$的各阶导数存在. 因此，所有的代数矩和绝对矩

$$\alpha_k=EX^k,\quad\beta_k=E|X|^k,\quad k=0,1,2,\cdots$$

存在，Φ_X有如下级数展开式：

$$\Phi_{X(z)}=\sum_{k=0}^{\infty}\frac{\mathrm{i}^k\alpha_k}{k!}z^k,\quad z\in\mathbf{C}.$$

且级数

$$\sum_{k=0}^{\infty}\beta_k\frac{z^k}{k!},\quad z\in\mathbf{C}$$

249

收敛.

要验证级数的收敛性，只需证明$\frac{\beta_{2k-1}}{(2k-1)!}(k\in\mathbf{N})$有下列形式的上界即可

$$\frac{\beta_{2k-1}}{(2k-1)!}\leqslant\frac{1}{2}\left(\frac{\alpha_{2k}}{(2k)!}(2k)+\frac{\alpha_{2k-2}}{(2k-2)!}\right).$$

由不等式

$$|x^{2k-1}| \leqslant \frac{1}{2}(x^{2k} + x^{2k-2}), \quad k = 1,2,\cdots,$$

知

$$\beta_{2k-1} = E|X^{2k-1}| \leqslant \frac{1}{2}(E(X^{2k})) + (E(X^{2k-2})) = \frac{1}{2}(\alpha_{2k} + \alpha_{2k-2}).$$

因此，对任意的 $A>0$，有

$$\begin{aligned}\sum_{k=0}^{\infty} \frac{\beta k}{k!}|y|^k &= \sum_{k=0}^{\infty} \frac{|y|^k}{k!} E|X|^k \\ &\geqslant \sum_{k=0}^{\infty} \frac{|y|^k}{k!} \int_{-A}^{A} |x|^k \mathrm{d}F_X(x) \\ &= \int_{-A}^{A} \sum_{k=0}^{\infty} \frac{|y|^k}{k!} |x|^k \mathrm{d}F_X(x) \\ &= \int_{-A}^{A} \mathrm{e}^{|xy|} \mathrm{d}F_X(x).\end{aligned}$$

所以，积分 $\int_{-\infty}^{\infty} \mathrm{e}^{|xy|} \mathrm{d}F_X(x)$ 存在，由此推出积分

$$\int_{-\infty}^{\infty} \mathrm{e}^{\mathrm{i}zx} \mathrm{d}F_X(x), \quad z \in \mathbf{C}$$

收敛，且当 z 为实数时等于 $\varphi(z)$. 那么对所有的 $z\in\mathbf{C}$，积分必等于 $\varphi(z)$，所以

$$E(\mathrm{e}^{\mathrm{i}zX}) = \mathrm{e}^{\mathrm{i}z\mu - \sigma^2 z^2/2}, \quad z \in \mathbf{C}$$

成立. ■

现在可以证明下列形式的 Girsanov 定理了.

定理 7.2.4(Girsanov) 设$\{B_t: t\in[0, T]\}$是概率 P 下的标准布朗运动. 如果 Q 定义为 $\mathrm{d}Q=L_T\mathrm{d}P$，其中 L_T 为(7.3)式所给定，则

$$B_t' = B_t + mt, \quad t \in [0,T],$$

[250] 是概率 Q 下的一个标准布朗运动.

证明 由 $B_t'-B_s'$ 独立于$\mathcal{F}_s$ 且 $A\in\mathcal{F}_s$ 知

$$\begin{aligned}E(\mathbf{1}_A \mathrm{e}^{\mathrm{i}u(B_t'-B_s')} L_t) &= E(\mathbf{1}_A \mathrm{e}^{\mathrm{i}u(B_t'-B_s')} L_t L_s^{-1} L_s) \\ &= E(\mathbf{1}_A L_s E(\mathrm{e}^{\mathrm{i}u(B_t'-B_s')} L_t L_s^{-1} | \mathcal{F}_s)) \\ &= E(\mathbf{1}_A L_s) E(\mathrm{e}^{\mathrm{i}u(B_t'-B_s')} L_t L_s^{-1}).\end{aligned}$$

现计算第二个因式

$$\begin{aligned}E(\mathrm{e}^{\mathrm{i}u(B_t'-B_s')} L_t L_s^{-1}) &= E\left(\exp\left\{\mathrm{i}u(B_t - B_s) + \mathrm{i}um(t-s) - m(B_t - B_s) - \frac{m^2}{2}(t-s)\right\}\right) \\ &= E\left(\exp\{(\mathrm{i}u - m)(B_t - B_s)\}\exp\left\{\mathrm{i}um(t-s) - \frac{m^2}{2}(t-s)\right\}\right).\end{aligned}$$

下面通过 B_t-B_s 的特征函数来计算第一个因式：因为 B_t-B_s 的特征函数为 $u \mapsto \mathrm{e}^{-u^2(t-s)/2}$，$\forall u\in\mathbf{C}$，所以

$$E\exp\{(\mathrm{i}u - m)(B_t - B_s)\} = E\exp\{\mathrm{i}(u + im)(B_t - B_s)\}$$

$$=\mathrm{e}^{-(u+\mathrm{i}m)^2(t-s)/2}.$$

因此，

$$\begin{aligned}E(\mathrm{e}^{\mathrm{i}u(B_t'-B_s')}L_tL_s^{-1})&=\exp\left\{-(u+\mathrm{i}m)^2(t-s)/2+\mathrm{i}um(t-s)-\frac{m^2}{2}(t-s)\right\}\\&=\exp\left\{(t-s)\left[-\frac{u^2}{2}-\mathrm{i}um+\frac{m^2}{2}+\mathrm{i}um-\frac{m^2}{2}\right]\right\}\\&=\exp\{-u^2(t-s)/2\}.\end{aligned}$$

定理得证. ■

7.3 等价鞅测度

本节的目的是确定一个等价鞅测度 P^*，使得贴现股票价格过程是一个 P^*-鞅. 因贴现股票价格过程为

$$S_{1t}^*=S_{1t}\mathrm{e}^{-rt}=S_{10}\exp\left((\mu-r)t-\frac{\sigma^2}{2}t+\sigma B_t\right),\tag{7.4}$$

为了确保几何布朗运动 $\exp\left(\sigma W_t-\frac{\sigma^2}{2}t\right)$继承了 W_t 的鞅性质，必须考虑右趋势项$-\frac{\sigma^2}{2}t$. 如果令

$$W_t=\frac{\mu-r}{\sigma}t+B_t,$$

251

则 S_{1t}^* 可以写成这种形式，但 W_t 在概率 P 下不是一个鞅. 由 Girsanov 定理知，存在一个等价的鞅测度 P^*，其定义为

$$P^*(A)=\int_A L_T\mathrm{d}P,\quad A\in\mathcal{F},$$

其中，

$$L_T=\exp\left(-\Theta B_T-\frac{\Theta^2}{2}T\right),\quad \Theta=\frac{\mu-r}{\sigma},$$

使得 W_t 在测度 P^* 下是一个鞅. 称 Θ 为**风险的市场价格**. 股票价格 S_{1t}为

$$S_{1t}=S_{10}\exp\left(\left(\mu-\frac{\sigma^2}{2}\right)t+\sigma B_t\right),$$

其中 B_t 是 P 下的标准布朗运动. 若令 $W_t=\frac{\mu-r}{\sigma}t+B_t$，则有

$$S_{1t}=S_{10}\exp\left(\left(r-\frac{\sigma^2}{2}\right)t+\sigma W_t\right),$$

其中 W_t 是 P^* 下的一个标准布朗运动. 也就是说为了计算 P^* 下相应的 S_{1t}，可以简单地用无风险利率 r 来代替 μ.

贴现价格过程 $S_{1t}^*=\exp\left(-\frac{\sigma^2}{2}t+\sigma W_t\right)$也可以写成如下形式：

$$S_{1t}^*=\int_0^t\sigma S_{1u}^*\mathrm{d}W_u,\quad t\in[0,T].$$

接下来，考虑贴现价值过程. 在 Itô积分 $\int_0^t \alpha_u \sigma S_{1u}^* \mathrm{d}W_u$ 有定义的前提下，有

$$V_t^* = V_0 + \int_0^t \alpha_u \sigma S_{1u}^* \mathrm{d}W_u, \quad t \in [0, T].$$

以后将交易策略限制为使该 Itô积分有定义的策略. 也就是说，以后只考虑满足

$$E^* \left(\int_0^T \left(\alpha_u S_{1u}^* \right)^2 \mathrm{d}u \right) < \infty$$

的自融资交易策略.

7.4 无套利定价与对冲

252

本节用前几节得到的结论来定价和套期保值 Black-Scholes 模型中的未定权益.

设 C：$(\Omega, \mathcal{F}, P) \to (\mathbf{R}, \mathcal{B})$是一个在时刻 T 提供报酬 C 的权益. 0 时刻的无套利价格为

$$\pi(C) = E^* (\mathrm{e}^{-rT} C) = \int \mathrm{e}^{-rT} C(\omega) \mathrm{d}P(\omega).$$

特别地，考虑欧式看涨期权 $C = (S_T - K)^+$，则需计算

$$E^* (\mathrm{e}^{-rT} C) = E^* (\mathrm{e}^{-rT} \max(S_T - K, 0)),$$

其中 $\log S_T$ 是在 P^* 下的对数正态分布，

$$\log S_T \overset{P^*}{\sim} \left(r - \frac{\sigma^2}{2} \right) T + \sigma W_T \sim N\left(\left(r - \frac{\sigma^2}{2} \right) T, \sigma^2 T \right).$$

计算的具体细节见 1.5.5 节.

考虑一个权益，即一个$\mathcal{F}_T$ 可测的随机变量，由它的$\mathcal{F}_T$ 可测性知，到期支付可能依赖于价格过程的全部路径，因此它是与路径有关的衍生品.

定义 7.4.1 一个权益称为**可复制的**(**可达的**)，如果存在一个适当的交易策略，$\varphi = \{\varphi_t : t \in [0, T]\}$(称为一个**对冲**)，使得

$$V_T(\varphi) = C.$$

假设给了一个权益 C 和一个可料过程$\{\alpha_t\}$，满足

$$C^* = \mathrm{e}^{-rT} C = \alpha_0 + \int_0^T \alpha_u \mathrm{d}S_{1u}^*.$$

那么可以通过选择 β_t 使得满足

$$V_t^*(\varphi) = \alpha_t S_{1t}^* + \beta_t \mathrm{e}^{-rt} = \alpha_0 + \int_0^t \alpha_u \mathrm{d}S_{1u}^*$$

的 $\varphi_t = (\alpha_t, \beta_t)$是一个自融资交易策略. 也就是说，对冲是通过银行账户融资实现的，对冲复制了权益 C，即 $V_T(\varphi) = C$.

设 P^* 是一个等价鞅测度，且 $E\int_0^T \alpha_t^2 \mathrm{d}t < \infty$，则 $\int_0^t \alpha_u \mathrm{d}S_{1u}^*$ 是一个 P^*-鞅，且

$$E^*(V_T^*(\varphi)) = \alpha_0.$$

从而 α_0 是无套利价格.

253

下面的定理提供了一种不需要明确的对冲策略就可以计算价值过程的方法.

定理 7.4.2 设 C_T 是一个满足 $E^*(C_T^2)<\infty$ 的权益. 若 C_T 是可复制的，则它在 t 时刻的价值为：

$$V_t = E^*(\mathrm{e}^{-r(T-t)}C_T|\mathcal{F}_t)$$

特别地，$V_0=E^*(\mathrm{e}^{-rT}C_T)$表示公平价格.

证明 设$\{\alpha_t\}$为一个复制对冲的股票数，满足

$$C_T^* = \alpha_0 + \int_0^T \alpha_u \mathrm{d}S_{1u}^*,$$

又设 $\varphi_t=(\alpha_t, \beta_t)$是相应的自融资策略，则由随机积分的性质知相应的价值过程 $V_t^* = \alpha_0 + \int_0^t \alpha_u \mathrm{d}S_{1u}^*, t\in[0,T]$ 是一个 P^*-鞅. 则

$$V_t^* = E^*(V_T^*|\mathcal{F}_t) = E^*(C_T^*|\mathcal{F}_t)$$

对所有的 $t\in[0, T]$均成立. 又因为 $C_T^*=\mathrm{e}^{-rT}C_T$，$V_t=\mathrm{e}^{-rt}V_t^*$，所以

$$V_t = E^*(\mathrm{e}^{-r(T-t)}C_T|\mathcal{F}_t), \quad t\in[0,T].$$

接下来的目标是证明 Black-Scholes 模型是完备的，即任一$\mathcal{F}_t$ 可测的权益都能被复制. 证明的基础是建立在下列连续时间的鞅表示定理之上的. ■

定理 7.4.3(表示定理) 设 $M=\{M_t:t\in[0, T]\}$是一个右连续的 P^*-鞅，则存在一个$(\mathcal{B}_{[0,T]}\otimes\mathcal{F}_T)$可测的适应过程 $\eta=\{\eta_t:t\in[0, T]\}$，使得

$$\int_0^T \eta_t^2 \mathrm{d}t < \infty$$

P^*-a. s. 成立，且

$$M_t = M_0 + \int_0^t \eta_s \mathrm{d}B_s^*, \quad t\in[0,T].$$

定理 7.4.4(完备性) 设 C 是一个$\mathcal{F}_t$ 可测的随机变量且满足 $E^*|C|<\infty$，则存在一个复制对冲 $\varphi=\{\varphi_t:t\in[0, T]\}$，即 $V_t(\varphi)=C$，且 $V_t^*(\varphi)=E^*(C|\mathcal{F}_t)$，$t\in[0, T]$.

证明 前面已经证明了任何复制对冲均满足：

$$V_t^* = E^*(C^*|\mathcal{F}_t), \quad t\in[0,T].$$

254

这表明等式的右边为贴现价值过程. 因此考虑 P^*-鞅：

$$M_t = E^*(C^*|\mathcal{F}_t), \quad t\in[0,T].$$

不妨设 M_t 是连续的. 表示定理保证了存在着满足 $\int_0^T \eta_t^2 \mathrm{d}t < \infty (P^*\text{-a. s.})$ 的适应过程$\{\eta_t\}$，使得

$$M_t = M_0 + \int_0^t \eta_s \mathrm{d}B_s^*.$$

为了确定一个自融资复制交易策略，作如下假设：如果 $\varphi_t=(\alpha_t, \beta_t)$是这样一个自融资对冲，则

$$V_t^* = V_0 + \int_0^t \alpha_u \mathrm{d}S_{1u}^*$$

成立. 因为

$$S_{1t}^* = \int_0^t \sigma S_{1u}^* dB_u^*$$

是一个 Itô过程，则有

$$\int_0^t \alpha_u dS_{1u}^* = \int_0^t \alpha_u \sigma S_{1u}^* dB_u^* .$$

由等式 $M_t - M_0 = \int_0^t \eta_u dB_u^*$ 和 $V_t^* - V_0 = \int_0^t \alpha_u \sigma S_{1u}^* dB_u^*$ ，解出 α_u，解得

$$\alpha_t = \frac{\eta_t}{\sigma S_{1t}^*}, \quad t \in [0,T].$$

令 $\beta_t = M_t - \alpha_t S_{1t}^*$，注意到 α_t 和 β_t 是适应的，则 $\varphi_t = (\alpha_t, \beta_t)$满足：

$$V_t^*(\varphi) = \alpha_t S_{1t}^* + \beta_t = M_t, \quad t \in [0,T].$$

接下来验证：φ_t 复制 C，并且是一个适当的自融资. 由

$$V_T^*(\varphi) = M_T = E^*(C^* \mid \mathcal{F}_T) = C^*$$

易得 φ_t 复制 C. 进一步，如果

255

$$V_t^*(\varphi) = V_0(\varphi) + \int_0^t \alpha_u dS_{1u}^*, t \in [0,T]$$

成立，则 φ_t 是自融资的. 首先验证 Itô积分 $\int_0^t \alpha_u dS_{1u}^*$ 是有定义的，为此，只需验证充分条件：

$$\int_0^T \alpha_u^2 du < \infty, P\text{-a. s.}$$

由于对任何 $t \in [0, T]$，均有 $S_{1t}^* > 0$，P^*-a. e.，则 $\inf_{t\in[0,T]} |S_{1t}^*| > 0$. 又因为 S_{1t}^* 是贴现股价过程，因此

$$\begin{aligned}\int_0^T \alpha_u^2 du &= \int_0^T \left(\frac{\eta_u}{\sigma S_{1u}^*}\right)^2 du \\ &\leqslant \left(\sigma \cdot \inf_{t\in[0,T]} |S_{1t}^*|\right)^{-2} \int_0^T \eta_u^2 du < \infty\end{aligned}$$

P^*-a. s. 成立 . 因为 $V_0(\varphi) = M_0$ 且 $\eta_u = \alpha_u \sigma S_{1u}^*$，则 φ_t 是自融资的，且满足

$$\begin{aligned}V_t^*(\varphi) &= M_t \\ &= V_0 + \int_0^t \eta_u dB_u^* \\ &= V_0 + \int_0^t \alpha_u \sigma S_{1u}^* dB_u^* \\ &= V_0 + \int_0^t \alpha_u dS_{1u}^*,\end{aligned}$$

由推论 7.1.5 知 φ_t 是自融资的. 最后，由

$$\int_0^T (\alpha_u S_{1u}^*)^2 du = \int_0^T \frac{\eta_u^2}{\sigma^2} du < \infty,$$

可得 φ_t 是适当的. ■

7.5 delta 对冲

设衍生品 C 的价值过程

$$V_t = \mathrm{e}^{rt} E^* (C\mathrm{e}^{-rT} \mid \mathcal{F}_t)$$

是关于 t 和 S_{1t} 的二阶连续可微函数，即对某个 C^2 函数 V：$[0, T]\times\mathbf{R}\to\mathbf{R}$，对 $(t, x)\in[0, T]\times\mathbf{R}$，$(t, x)\mapsto V(t, x)$，有

$$V_t = V(t, S_{1t}), \quad t \in [0, T].$$

256

记 V 的偏导数为 $\frac{\partial V}{\partial t}$ 和 $\frac{\partial V}{\partial x}$ 等. 设 $\{\varphi_t\}$ 是一个自融资策略，则由定义 7.1.3，有

$$\begin{aligned} V_t(\varphi) &= V_0(\varphi) + \int_0^t \alpha_u \mathrm{d}S_{1u} + \int_0^t \beta_u \mathrm{d}S_{0u} \\ &= V_0(\varphi) + \int_0^t \alpha_u \mathrm{d}S_{1u} + \int_0^t \beta_u r \mathrm{e}^{ru} \mathrm{d}u. \end{aligned}$$

对 $V(t, S_{1t})$ 使用 Itô公式得

$$\begin{aligned} V_t = V_0 &+ \int_0^t \left[\frac{\partial V}{\partial t}(u, S_{1u}) + \frac{\partial V}{\partial x}(u, S_{1u}) \mu S_{1u} + \frac{1}{2} \frac{\partial^2 V}{\partial x^2}(u, S_{1u}) \sigma^2 S_{1u}^2 \right] \mathrm{d}u \\ &+ \int_0^t \frac{\partial V}{\partial x}(u, S_{1u}) \sigma S_{1u} \mathrm{d}B_u. \end{aligned}$$

因为 $\mathrm{d}S_{1t} = \mu S_{1t}\mathrm{d}t + \sigma S_{1t}\mathrm{d}B_t$，上式可改写为

$$V_t = V_0 + \int_0^t \frac{\partial V}{\partial x}(u, S_{1u}) \mathrm{d}S_{1u} + \int_0^t \left[\frac{\partial V}{\partial t}(u, S_{1u}) + \frac{1}{2} \frac{\partial^2 V}{\partial x^2}(u, S_{1u}) \sigma^2 S_{1u}^2 \right] \mathrm{d}u.$$

令被积函数相等即可得到 **delta 对冲**

$$\begin{aligned} \alpha_t &= \frac{\partial V}{\partial x}(t, S_{1t}), \\ \beta_t &= \left[\frac{\partial V}{\partial t}(t, S_{1t}) + \frac{1}{2} \frac{\partial^2 V}{\partial x^2}(t, S_{1t}) \sigma^2 S_{1t}^2 \right] / (r\mathrm{e}^{rt}). \end{aligned}$$

将它们代入 $V_t = \alpha_t S_{1t} + \beta_t S_{0t}$ 就得到

$$rV_t = r \frac{\partial V}{\partial x}(t, S_{1t}) S_{1t} + \frac{\partial V}{\partial t}(t, S_{1t}) + \frac{1}{2} \frac{\partial^2 V}{\partial x^2}(t, S_{1t}) \sigma^2 S_{1t}^2.$$

这表明价值过程满足一个偏微分方程.

7.6 与时间有关的波动率

迄今为止，所有股价的波动率都是假定为一个非随机的常量. 然而，来自金融市场的实际数据表明这是一个不真实的假设. 本节假定波动率是 t 的函数. 以随机微分方程为出发点，假设股价 S_{1t} 满足模型

$$\mathrm{d}S_{1t} = \mu_t S_{1t} \mathrm{d}t + \sigma_t S_{1t} \mathrm{d}B_t,$$

257

其中 σ：$[0, T]\to[0, \infty]$ 是满足

$$\int_0^T \sigma_u^2 \mathrm{d}u < \infty$$

的某个确定性函数，这就确保了积分 $\int_0^t \sigma_u \mathrm{d}B_u, t \in [0,T]$ 是有定义的，再假定适应的漂移过程 μ_t 满足

$$\int_0^T |\mu t| \mathrm{d}t < \infty$$

a.s. 成立.

解上面的方程意味着要求如下积分方程的解

$$S_{1t} = \int_0^t \mu S_{1u} \mathrm{d}u + \int_0^t \sigma_t \mathrm{d}B_t, \quad t \in [0,T].$$

在例 6.6.9 中已经指出广义布朗运动

$$S_{1t} = S_{10} \exp\left(\int_0^t \left(\mu - \frac{\sigma_u^2}{2}\right) \mathrm{d}u + \int_0^t \sigma_u \mathrm{d}B_u\right),$$

$t \in [0, T]$，是该方程的解，其中 S_{10} 为初始价值，$\int_0^t \sigma_u \mathrm{d}B_u, t \in [0,T]$ 是一个 P 鞅. 注意到 S_{1t} 也可以写成

$$S_{1t} = S_{10} \exp\left(\left(\mu - \frac{\overline{\sigma^2}(0,t)}{2}\right) t + \int_0^t \sigma_u \mathrm{d}B_u\right),$$

其中，对 $0 \leqslant s \leqslant t \leqslant T$，

$$\overline{\sigma^2}(s,t) = \frac{1}{t-s} \int_s^t \sigma_u^2 \mathrm{d}u.$$

则贴现价格过程为：

$$S_t^* = S_{10} \exp\left(\left(\mu - r - \frac{\overline{\sigma^2}(0,t)}{2}\right) t + \int_0^t \sigma_u \mathrm{d}B_u\right),$$

这与常数波动率下的形式是一样的，对比(7.4)式，σ^2 被 $\overline{\sigma^2}(0, t)$ 代替，σB_u 被 $\int_0^t \sigma_u \mathrm{d}B_u$ 代替. 所以这种模型的欧式看涨期权的价格同样由 Black-Scholes 公式给出，只需将 σ 替代成 $\sqrt{\overline{\sigma^2}(0, T)}$ 即可.

函数 σ_t 可由如下的隐含波动率推出. 令一个固定的敲定价格和到期时间 T 的 Black-Scholes 公式得到的价格和实际期权价格相等，可以确定到期时间为 T 的期权在 t 时刻的隐含波动率 $\sigma_{\mathrm{imp}}(T)$. 那么对等式

$$\frac{1}{T-t} \int_t^T \sigma_u^2 \mathrm{d}u = \sigma_{\mathrm{imp}}^2(T)$$

两边求导可得

258

$$\sigma_T^2 = 2\sigma_{\mathrm{imp}}^2(T)(\sigma_{\mathrm{imp}}^2)'(T)(T-t) + \sigma_{\mathrm{imp}}^2(T),$$

对任何到期时间 T 均成立. 因此得到了依赖于时间的波动率的函数形式.

7.7 Black-Scholes 模型的推广

本节假定波动率 σ_t 是一个适应过程，这比假定它是一个确定性的函数更符合实际. 同样以随机微分方程为出发点，设股价 S_{1t} 满足如下模型

$$dS_{1t} = \mu_t S_{1t} dt + \sigma_t S_{1t} dB_t, \tag{7.5}$$

其中 σ：$[0, T] \to [0, \infty]$是某个(可能随机但适应的)函数，且满足：

$$\int_0^T \sigma_u^2 du < \infty,$$

a. s. 成立. 这保证了积分 $\int_0^t \sigma_u dB_u, t \in [0, T]$ 是有定义的. 且假定适应的漂移过程 μ_t 满足

$$\int_0^T |\mu_t| dt < \infty,$$

a. s. 成立. 回顾一下之前对随机微分方程(7.5)的推导过程：从 t 到 $t + dt$，其中 dt 很小，股价按线性漂移项 $\mu_t S_{1t} dt$(其在时刻 t 是已知的)变化，并受随机干扰项 $\sigma_t S_{1t} \sqrt{dt} Z$ 的影响，它与 S_{1t}成比例，其中 $Z \sim N(0, 1)$. 换句话说，(7.5)式将相对变化 dS_{1t}/S_{1t}刻画为：

$$\frac{dS_{1t}}{S_{1t}} \approx \mu_t dt + \sqrt{dt} \sigma_t Z,$$

其中 μ_t 和 σ_t 在时刻 t 是已知的.

银行账户(货币市场)可以由一个带无风险利率(短期利率，即时利率)r_t 的局部无风险债券表示，即

$$S_{0t} = r_t dt, \quad t \in [0, T],$$

其中

$$\int_0^T |r_u| du < \infty,$$

几乎必然成立.

上述方程的具体解法如下：由例 6.6.9 的结论知广义布朗运动

$$S_{1t} = S_{10} \exp\left(\int_0^t \left(\mu_u - \frac{\sigma_u^2}{2}\right) du + \int_0^t \sigma_u dB_u\right), \quad t \in [0, T]$$

259

是上述随机微分方程的解，其中 S_{10} 表示在 0 时刻的初始股价. 对银行账户有如下的显式解：

$$S_{0t} = S_{00} \exp\left(\int_0^t r_u du\right), \quad t \in [0, T].$$

其中不妨假定 $S_{00} = 1$.

为了更深入的讨论，需要下列一般形式的 Girsanov 定理.

定理 7.7.1(Girsanov 定理) 设$\{B_t\}$是实际概率测度 P 下的布朗运动，$\{\gamma_t\}$是$\mathcal{F}_t$ 适应过程，且满足 Novikov 条件

$$E\left[\exp\left(\frac{1}{2}\int_0^T \gamma_t^2 dt\right)\right] < \infty.$$

定义

$$L_t = \exp\left(-\int_0^t \gamma_u dB_u - \frac{1}{2}\int_0^t \gamma_u^2 du\right),$$

设 P^* 是按下列方式定义的概率测度

$$P^*(A)=\int_A L_T(\omega)\mathrm{d}P(\omega),\quad t\in[0,T].$$

那么，在概率测度 P^* 下，过程

$$B_t^*=B_t+\int_0^t\gamma_u\mathrm{d}u,\quad t\in[0,T]$$

是一个标准的布朗运动.

下面确定一个等价鞅测度，使得贴现股价在该测度下是一个鞅．为此，设

$$B_t^*=B_t+\int_0^t\gamma_u\mathrm{d}u,\quad t\in[0,T].$$

则贴现股价过程为

$$\begin{aligned}S_{1t}^*&=S_{1t}\exp\left(-\int_0^t r_u\mathrm{d}u\right)\\&=S_{10}\exp\left\{\int_0^t(\mu_u-r_u-\frac{\sigma_u^2}{2})\mathrm{d}u+\int_0^t\sigma_u\mathrm{d}B_u\right\},\end{aligned}$$

它是如下随机微分方程的解：

260

$$\mathrm{d}S_{1t}^*=(\mu_t-r_t)S_{1t}^*\mathrm{d}t+\sigma_tS_{1t}^*\mathrm{d}B_t.$$

即

$$S_{1t}^*=\int_0^t(\mu_u-r_u)S_{1u}^*\mathrm{d}u+\int_0^t\sigma_uS_{1u}^*\mathrm{d}B_u.$$

注意到 B_t^* 是一个漂移为 γ_t，波动为 1 的 Itô过程，由定理 6. 6. 5 的积分法则知

$$\int_0^t\sigma_uS_{1u}^*\mathrm{d}B_u^*=\int_0^t\sigma_u\gamma_uS_{1u}^*\mathrm{d}u+\int_0^t\sigma_uS_{1u}^*\mathrm{d}B_u.$$

因此，用 B_t^* 表示 S_{1t}^* 可得：

$$S_{1t}^*=\int_0^t(\mu_u-r_u-\gamma_u\sigma_u)S_{1u}^*\mathrm{d}u+\int_0^t\sigma_uS_{1u}^*\mathrm{d}B_t^*.$$

若令 $\gamma_t=\frac{\mu_t-r_t}{\sigma_t}$，$t\in[0,\ T]$，则可消去右边第一项，此时如果 B_t^* 是鞅，则 S_{1t}^* 也是一个鞅．对于这样定义的 γ_t，由 Girsanov 定理知，概率测度 P^* 就是所需的概率测度．称 γ_t 为风险市场价格.

显然，在 P^* 下，未贴现的股价 S_{1t} 为

$$S_{1t}=S_{10}+\int_0^t\mu_uS_u\mathrm{d}u+\int_0^t\sigma_uS_u\mathrm{d}B_t,$$

其平均回报率等于无风险利率．事实上，将 B_t 换为 B_t^* 得到

$$\int_0^t\sigma_uS_u\mathrm{d}B_t=\int_0^t\sigma_uS_{1u}\mathrm{d}B_u^*-\int_0^t\sigma_u\gamma_uS_{1u}\mathrm{d}u,$$

所以

$$\begin{aligned}S_{1t}&=S_{10}+\int_0^t(\mu_u+\sigma_u\gamma_u)S_u\mathrm{d}u+\int_0^t\sigma_uS_u\mathrm{d}B_t\\&=S_{10}+\int_0^t r_uS_u\mathrm{d}u+\int_0^t\sigma_uS_u\mathrm{d}B_t.\end{aligned}$$

注意，将模型由实际概率 P 下的模变化为风险中性测度 P^* 下的模型，只改变了漂移，波动没有改变.

7.8 评注与延伸阅读

在一定程度上，本书的内容受到了 Williams(2006)的启发，更深入的学习可阅读 Etheridge(2002)和 Shiryaev(1999). 平方可积鞅的表示定理的一本经典参考文献是 Kunita 和 Watanabe(1967)；读者也可以参考(Karatzas 和 Shreve，1991，第 3 章，定理 4.15). 进一步的讨论可在 Jacob 和 Shiryaev(2003)中找到. 建立在解析特征函数基础上的 Girsanov 定理的初等证明是选自 Nualart(2011)，一般情形(定理 7.7.1)的证明可以在很多文献中找到，比如 Revuz 和 Yor(1999). 要了解(解析)特征函数的更多内容，读者可参阅 Lukacs(1970). Shreve(2004)也较好地推导了 Girsanov 定理的一般情形. 期权价格与偏微分方程的关系在很多著作中都有讨论，比如 Shreve(2004).

参考文献

Etheridge A. (2002) *A Course in Financial Calculus*. Cambridge University Press, Cambridge.

Jacod J. and Shiryaev A.N. (2003) *Limit Theorems for Stochastic Processes*. vol. 288 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]* 2nd edn. Springer-Verlag, Berlin.

Karatzas I. and Shreve S.E. (1991) *Brownian Motion and Stochastic Calculus*. vol. 113 of *Graduate Texts in Mathematics* second edn. Springer-Verlag, New York.

Kunita H. and Watanabe S. (1967) On square integrable martingales. *Nagoya Math. J.* **30**, 209–245.

Lukacs E. (1970) *Characteristic Functions*. Hafner Publishing Co., New York. 2nd edition, revised and enlarged.

Nualart D. (2011) *Stochastic Calculus*. Lecture Notes.

Revuz D. and Yor M. (1999) *Continuous Martingales and Brownian Motion*. vol. 293 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]* 3rd edn. Springer-Verlag, Berlin.

Shiryaev A.N. (1999) *Essentials of Stochastic Finance: Facts, Models, Theory*. vol. 3 of *Advanced Series on Statistical Science & Applied Probability*. World Scientific Publishing Co. Inc., River Edge, NJ. Translated from the Russian manuscript by N. Kruzhilin.

Shreve S.E. (2004) *Stochastic Calculus for Finance. II: Continuous-time Models*. Springer Finance. Springer-Verlag, New York.

Williams R.J. (2006) *Introduction to the Mathematics of Finance*. vol. 72 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI.

261 ~ 262

第8章　离散时间过程的极限理论

随机序列的极限理论在金融数据分析中起着重要的作用，因为在很多时候，生成过程的原始数据不足以形成适合的参数模型．另外，对参数模型进行估计和推断分析也要用到非参或无模型的概念，以及渐近分布理论的一些结果．

本章从判别构建估计方法的平均过程的相关时间序列极限分布的大数定理开始，其后的许多结论直接或间接地依赖于这一结果．大数定律指出：当 $T\to\infty$ 时，

$$\overline{X}_T=\frac{1}{T}\sum_{t=1}^{T}X_t\xrightarrow{P}E(X_1).$$

即使是相关的时间序列，中心极限定理也给出了缩放误差 $\sqrt{T}(\overline{X}_T-E(X_1))$ 的近似分布．一个更一般的中心极限定理指出：存在某个 $\eta\in(0,\infty)$，使得

$$\sqrt{T}(\overline{X}_T-E(X_1))\xrightarrow{d}N(0,\eta^2).$$

上述两个结果确保了大数定理和中心极限定理这两个基本概率结论对弱正则性条件下的相依序列也是适用的．

回归是统计学中数据分析的一个基本工具．在金融应用中，自变量通常是随机的．因此，本章讨论一般条件下带随机自变量的多元线性回归，它包含了一个时间序列关于它的滞后值的回归．本章引入了几个有用的极限定理，并将它们应用于鞅差序列．

263

本章的下一内容是介绍非参数密度估计和非参数回归分析．在过去的很长一段时间里，非参数方法在金融实证分析中起着次要的作用．但是现在情况已经极大地改变了，非参数方法变得无处不在．因此，本章将对一些主要结果给出详细的介绍．

线性过程为研究 ARMA(p,q)过程提供了正确的框架．Beveridge-Nelsom 分解为用初等方法推导线性过程的中心极限定理提供了一个巧妙的工具．因此，本章专门用一小节来介绍这一漂亮且有用的结果．

本章还将讨论 α 混合过程．尽管相关计算的推导很少涉及，但是 α 混合系数的依赖度的计算是非常直观的，且在不需要假设基本的时间序列是一个线性过程甚至属于一个参数类的情况下可以得出一些强有力的结果．

本章最后讨论样本自协方差的渐近性及长期方差参数的 Newey-West 估计量，它在渐近公式中经常出现．

8.1　相关时间序列的极限定理

由于鞅差序列是不相关的，像 ARMA(p,q)过程这类时间序列关于滞后长度 $|h|>1$ 的自协方差不为0．但它们的算术平均仍然收敛于它们的均值，由此得到当滞后增加时，自协方差的衰减非常迅速．

为了理解相关性对概率计算的影响(通常是大到不可忽视的地步)，考虑下列平稳高斯时间序列

$$\mu = E(X_t), \quad \sigma^2 = \text{Var}(X_t).$$

自协方差为

$$\gamma(h) = E(X_t - \mu)(X_{t+|h|} - \mu), \quad h \in \mathbf{Z}.$$

假设可以观测过程到时刻 T 为止的值，其样本为 X_1，…，X_T. 未知期望 μ 的估计量的一个自然表达式为

$$\hat{\mu}_T = \frac{1}{T}\sum_{t=1}^{T} X_t.$$

由平稳性，$E(\hat{\mu}_T) = \mu$. 那么方差又如何估计呢？被加项 X_t 不是独立的，所以计算 Var$(\hat{\mu}_T)$的表达式为

$$\begin{aligned}\text{Var}(\hat{\mu}_T) &= E(\hat{\mu}_T - \mu)^2 \\ &= \frac{1}{T^2}\sum_{t=1}^{T}\sum_{s=1}^{T}\text{Cov}(X_s, X_t).\end{aligned}$$

264

式中双重求和把 $T \times T$ 维对称矩阵的所有元素 $\text{Cov}(X_s, X_t)$全部加起来. 换成将所有对角线元素相加得到

$$\begin{aligned}\text{Var}(\hat{\mu}_T) &= \frac{1}{T^2}\left\{2\sum_{k=1}^{T-1}(T-k)\gamma(k) + T\gamma(0)\right\} \\ &= \frac{1}{T}\sum_{|k|<T}\left(1 - \frac{|k|}{T}\right)\gamma(k).\end{aligned}$$

等价的表达式为

$$\eta_T^2 = \text{Var}(\hat{\mu}_T) = \frac{1}{T}\text{Var}(X_1) + \frac{2}{T^2}\sum_{k=1}^{T-1}(T-k)\gamma(k).$$

这是一个精确的计算公式，由此得到下列的精确分布结论.

命题 8.1.1 假设$\{X_t\}$是一个高斯时间序列，自协方差为 $\gamma(h)$，则对所有的 $T \in \mathbf{N}$，有

$$\sqrt{T}(\hat{\mu}_T - \mu) \sim N(0, \eta_T^2),$$

其中

$$\eta_T^2 = \sum_{|k|<T}\left(1 - \frac{|k|}{T}\right)\gamma(k).$$

上述命题有一个直接的推论：$\hat{\mu}_T = \frac{1}{T}\sum_{t=1}^{T} X_t$ 依概率收敛于它的期望μ，即弱大数定律

$$\hat{\mu}_T \xrightarrow{P} \mu,$$

当 $T \to \infty$时成立，且收敛速度为 $T^{1/2}$. 现有下列两个问题：

(ⅰ) 当时间序列是非高斯过程时，收敛还会成立吗？

(ⅱ) 在相关性满足适当的条件下，$\sqrt{T}(\hat{\mu}_T - \mu)$会收敛于一个正态分布吗？

第一个问题的回答是简单的：在滞后增大时，自协方差趋于 0 的弱假设下，收敛仍成立. 但是中心极限定理成立会牵涉到更多的条件，它需要在时间序列的相关性上有更强的

假定. 接下来将介绍几个不同条件下的中心极限定理，这些定理对于处理带离散时间的相关过程的模型是非常有用的.

265

回忆 Cesaro 和的性质，如果$\{\alpha_n: n\in \mathbf{N}\}$是一个实数列，满足当$n\to\infty$时，$\alpha_n\to\alpha\in\mathbf{R}$，那么当$n\to\infty$时，Cesaro 平均$\frac{1}{n}\sum_{i=1}^{n}\alpha_i$也收敛于$\alpha$. 记

$$\overline{X}_T=\frac{1}{T}\sum_{t=1}^{T}X_t,\quad T\geqslant 1.$$

定理 8.1.2(大数定律) 设$\{X_t\}$是一个弱平稳的时间序列，且均值$\mu=E(X_1)$，自协方差为$\gamma(k)$，$k\in\mathbf{Z}$.

(i) 若$\lim\limits_{k\to\infty}\gamma(k)=0$，则

$$\mathrm{Var}(\overline{X}_T)\to 0,\quad T\to\infty,$$

且当$T\to\infty$时，

$$\overline{X}_T\xrightarrow{P,L_2}\mu.$$

(ii) 若

$$\sum_{k\in\mathbf{Z}}|\gamma(k)|<\infty,$$

则当$T\to\infty$时，

$$\mathrm{Var}(\sqrt{T}\,\overline{X}_T)=TE(\overline{X}_T-\mu)^2\to\sum_{k\in\mathbf{Z}}\gamma(k).$$

证明 (i)注意到:

$$\gamma(k)-\frac{|k|}{T}\gamma(k)\leqslant\max\{0,\gamma(k)\}.$$

因此，当$T\to\infty$时，Cesaro 平均$\frac{1}{T}\sum_{k=1}^{T}\left(1-\frac{|k|}{T}\right)\gamma(k)$收敛于 0，由此可得

$$\lim_{T\to\infty}\mathrm{Var}(\overline{X}_T)=0.$$

由切比雪夫不等式知，对任意$\varepsilon>0$，当$T\to\infty$时，有

266

$$P(|\overline{X}_T-\mu|>\varepsilon)\leqslant\frac{\mathrm{Var}(\overline{X}_T)}{\varepsilon^2}\to 0.$$

(ii) 考虑

$$T\mathrm{Var}(\overline{X}_T)=\sum_{|k|<T}\left(1-\frac{|k|}{T}\right)\gamma(k)=\int f_T(k)\mathrm{d}\upsilon(k),$$

其中υ为$\mathbf{Z}$上的计数测度，且

$$f_T(k)=\left(1-\frac{|k|}{T}\right)\gamma(k)\mathbf{1}(|k|<T),\quad k\in\mathbf{Z}.$$

f_T的控制函数为$f(k)=|\gamma(k)|$，$k\in\mathbf{Z}$，由假设知$f(k)$是υ可积的，因为$\int|f|\mathrm{d}\nu=\sum_{k\in\mathbf{Z}}|\gamma(k)|<\infty$. 进一步，$f$是逐点收敛的. 事实上，对任意给定的$k\in\mathbf{Z}$，

$$f_T(k)=\left(1-\frac{|k|}{T}\right)\gamma(k)\mathbf{1}(|k|<T)\to\gamma(k)=:f(k).$$

由控制收敛定理知：当 $T\to\infty$ 时，有

$$T\mathrm{Var}(\overline{X}_T)=\int f_T(k)\mathrm{d}v(k)\to\int f(k)\mathrm{d}v(k)=\sum_{k\in\mathbf{Z}}|\gamma(k)|,$$

得证. ■

评注 8.1.3 结论(ⅰ)表明 $\mathrm{Var}(\overline{X_T})$ 收敛于 0，这就得到了弱大数定律. 定理 8.1.2 的结论(ⅱ)则更强，它证明了 $T\mathrm{Var}(\overline{X_T})$(或 $\mathrm{Var}(\sqrt{T}\,\overline{X}_T)$)的收敛性，这就确保了统计量 $\sqrt{T}(\hat{\mu}_T-\mu)$ 的渐近方差的存在性，人们希望该统计量在一定的相关性条件下对大样本是正态分布的.

η_T^2 的极限

$$\eta^2=\sum_{h\in\mathbf{Z}}\gamma(h)=\gamma(0)+2\sum_{h=0}^{\infty}\gamma(h) \tag{8.1}$$

是缩放样本均值的渐近方差，称其为**长期方差**. 下面简要介绍一个例子.

▶**例 8.1.4** 假设 $\{X_t\}$ 是一个 AR(1) 过程，其 AR 参数 ρ 满足 $|\rho|<1$，即

$$X_t-\mu=\rho(X_{t-1}-\mu)+\varepsilon_t,$$

其扰动项是独立同分布的，且服从正态分布 $N(0,\sigma^2)$. 因为

$$\gamma(h)=\rho^{|h|}\frac{\sigma^2}{1-\rho^2},$$

267

所以长期方差为下列式子的极限：

$$\eta_T^2=\frac{\sigma^2}{1-\rho^2}+2\frac{\sigma^2}{1-\rho^2}\sum_{h=1}^{T-1}\rho^h=\frac{\sigma^2}{1-\rho^2}+2\frac{\sigma^2}{1-\rho^2}\frac{1-\rho^T}{1-\rho}.$$

上述大数定律表明时间平均

$$\overline{X}_T(\omega)=\frac{1}{T}\sum_{t=1}^{T}X_t(\omega)$$

收敛于其期望，也就是说，对所有 ω，(积分)加权平均为

$$EX_t=\int X_t(\omega)\mathrm{d}P(\omega).$$

这样的结论称为**遍历定理**，称满足遍历定理的过程 $\{X_t\}$ 为**遍历的**. 弱大数定律可以弱化为下列形式. ◀

定理 8.1.5(遍历定理) 任何弱平稳过程 $\{X_t\}$ 都是遍历的，即当 $T\to\infty$ 时，有

$$\overline{X}_T\xrightarrow{P}E(X_1).$$

对一个严平稳过程 $\{X_t,\ t\in\mathbf{Z}\}$，设 $E(X_1)=0$，考虑自然滤子 $\mathcal{F}_t=\sigma(X_s:\ s\leqslant t)$. 回顾一下：条件期望 $E(X_t|\mathcal{F}_s)(s\leqslant t)$ 是 X_t 和 $\|E(X_t|\mathcal{F}_s)\|_2$ 的 L_2 最优预测，$\|E(X_t\mid F_s)\|_2$ 随 $|s-t|$ 的增大而减小，它是用 X_s，X_{s-1}，…的可测函数来预测 X_t 的准确程度的度量. 作为一个例子，考虑下列线性过程

$$X_t=\sum_{i=0}^{\infty}\theta_i\varepsilon_{t-i},\quad t\in\mathbf{Z},$$

其误差项$\{\varepsilon_t\}$是独立同分布的，且服从正态分布$N(0, \sigma^2)$，$\sigma>0$. 给定X_s，X_{s-1}，…，则X_t的最优预测为

$$E(X_t \mid \mathcal{F}_s) = \sum_{i=t-s}^{\infty} \theta_i \varepsilon_{t-i}.$$

当$t-s \to \infty$时，其范数

$$\| E(X_t \mid \mathcal{F}_s) \|_2 = \sigma \sqrt{\sum_{i=t-s}^{\infty} \theta_i^2}$$

收敛于0. 下列广义中心极限定理证实了$\frac{1}{\sqrt{T}} \sum_{t=1}^{T} X_t$的渐近正态性，只要$E(X_t \mid \mathcal{F}_s) \|_2 (s=t-1, t-2, \cdots)$足够快地收敛于0，就能得到该级数收敛. 对线性过程来说，如果时间序列模型是由独立同分布的扰动过程导出的，那么这个条件是比较容易验证的. 该中心极限定理也可用来证明混合过程的中心极限定理及推导样本自协方差函数.

268

作为准备，需要引入下列引理来建立条件期望$E(X \mid Y)$的L_2范数与协方差$\mathrm{Cov}(X, Y)$的关系.

引理 8.1.6 假设X是一个随机变量，满足$E(X)=0$，$\mathcal{A}$是一个σ代数，则

$$\| E(X \mid \mathcal{A}) \|_{L_2} = \sup\{E(XY): Y \text{ 是}\mathcal{A}\text{可测的}, \|Y\|_{L_2} = 1\}.$$

证明 令Y为$\mathcal{A}$可测的随机变量，满足$\|Y\|_{L_2}=1$，由柯西-施瓦兹不等式知，

$$\begin{aligned} E(XY) &\leqslant E(E(XY \mid \mathcal{A})) \\ &= E(YE(X \mid \mathcal{A})) \\ &\leqslant \|Y\|_{L_2} \| E(X \mid \mathcal{A}) \|_{L_2} \\ &= \| E(X \mid \mathcal{A}) \|_{L_2} \end{aligned}$$

等式成立当且仅当Y和$E(X \mid \mathcal{A})$是线性相关的，此时有

$$Y = \frac{E(X \mid \mathcal{A})}{\| E(X \mid \mathcal{A}) \|_{L_2}}.$$

■

定理 8.1.7 令$\{X_t\}$是一个严平稳序列，满足$E(X_1)=0$，且满足遍历性定理. 如果

$$\sum_{n=1}^{\infty} \| E(X_0 \mid \mathcal{F}_{-n}) \|_2 < \infty,$$

那么

$$\frac{1}{\sqrt{T}} \sum_{t=1}^{T} X_t \xrightarrow{d} N(0, \sigma^2), \quad T \to \infty,$$

其中

$$\sigma^2 = E(X_1^2) + 2 \sum_{h=1}^{\infty} E(X_1 X_{1+h})$$

绝对收敛.

证明 定义

$$Z_t = \sum_{j=0}^{\infty} E(X_{t+j} \mid \mathcal{F}_{t-1}),$$

$$Y_t = \sum_{j=0}^{\infty}[E(X_{t+j} | \mathcal{F}_t) - E(X_{t+j} | \mathcal{F}_{t-1})].$$

269

下面证明：

（ⅰ）上述序列 Z_t 和 Y_t 是 L_2 收敛的.

（ⅱ）下列分解

$$X_t = Y_t + Z_t - Z_{t+1}$$

对所有 t 均 a. s. 成立.

（ⅲ）当 $T\to\infty$时，有

$$\frac{Z_1 - Z_{T+1}}{\sqrt{T}} \xrightarrow{P} 0.$$

（ⅳ）$\{Y_t\}$是一个 L_2 鞅差序列，从而它满足定理 8.3.8 的条件.

（ⅴ）级数 σ^2 是绝对收敛的.

为了证明（ⅰ），只需要证明

$$\sum_{j=0}^{\infty}\| E(X_{t+j} | \mathcal{F}_{t-1})\|_2 < \infty.$$

因为$\{X_t\}$是平稳的，由引理 8.1.6 知，对 $j=0$，1，…有

$$\begin{aligned}\| E(X_{t+j} | \mathcal{F}_{t-1})\|_2 &= \sup\{E(X_{t+j}Z) : Z \text{ 是}\mathcal{F}_{t-1} \text{ 可测的}, \|Z\|_2 = 1\}\\ &= \sup\{E(X_{t+j}f(X_{t-1},\cdots)) : f \text{ 是可测的}, \|f(X_{t-1},\cdots)\|_2 = 1\}\\ &= \sup\{E(X_0 f(X_{-j-1},\cdots)) : f \text{ 是可测的}, \|f(X_{-j-1},\cdots)\|_2 = 1\}\\ &= \|E(X_0 | \mathcal{F}_{-j-1})\|_2,\end{aligned}$$

所以，由假定有

$$\begin{aligned}\sum_{j=0}^{\infty}\|E(X_{t+j} | \mathcal{F}_{t-1})\|_2 &= \sum_{j=0}^{\infty}\|E(X_0 | \mathcal{F}_{-j-1})\|_2\\ &= \sum_{j=1}^{\infty}\|E(X_0 | \mathcal{F}_{-j})\|_2 < \infty.\end{aligned}$$

270

下面证明（ⅱ），由 Y_t 和 Z_t 的定义，$E(X_{t+j} | \mathcal{F}_{t-1})$的项在相加时抵消了，所以

$$\begin{aligned}Z_t + Y_t &= \sum_{j=0}^{\infty}E(X_{t+j} | \mathcal{F}_t)\\ &= E(X_t | \mathcal{F}_t) + \sum_{j=1}^{\infty}E(X_{t+j} | \mathcal{F}_t)\\ &= X_t + \sum_{j=0}^{\infty}E(X_{t+j+1} | \mathcal{F}_t)\\ &= X_t + Z_{t+1}, \quad \text{a. s.}\end{aligned}$$

现在验证（ⅲ），首先需要证明$\{Z_t\}$是严平稳的. 这需要一定的技巧，下面是证明的具体步骤. 证明是建立在 $E(X_s | \mathcal{F}_t)$是移位不变之上的. 显然，存在可测函数 f 使得

$$E(X_s | \mathcal{F}_t) = f(X_t, X_{t-1}, \cdots).$$

下面证明对任意 $h\in\mathbf{Z}$，$f(X_{t+h}, X_{t+h-1}, ..)$是 $E(X_{s+h} | \mathcal{F}_{t+h})$的修正. 即对所有的 $A\in$

$\mathcal{F}_{t+h}$，有

$$\int_A f(X_{t+h}, X_{t+h-1}, \cdots)\mathrm{d}P = \int_A X_{s+h}\mathrm{d}P.$$

取 $A=\{(X_{t+h}, X_{t+h-1}, ..)\in B\}$，其中可测集 $B\subset\mathbf{R}^{\mathbf{N}_0}$，则有

$$\begin{aligned}\int_A f(X_{t+h}, X_{t+h-1}, \cdots)\mathrm{d}P &= \int_{\{X_{t+h}, X_{t+h-1}, \cdots)\in B\}} f(X_{t+h}, X_{t+h-1}, \cdots)\mathrm{d}P \\ &= \int_{\{(X_t, X_{t-1}, \cdots)\in B\}} f(X_t, X_{t-1}, \cdots)\mathrm{d}P \\ &= \int_{\{(X_t, X_{t-1}, \cdots)\in B\}} E(X_s|\mathcal{F}_t)\mathrm{d}P \\ &= \int_{\{(X_t, X_{t-1}, \cdots)\in B\}} X_s\mathrm{d}P \\ &= \int_{\{(X_{t+h}, X_{t+h-1}, \cdots)\in B\}} X_{s+h}\mathrm{d}P \\ &= \int_A X_{s+h}\mathrm{d}P.\end{aligned}$$

由此可得

271

$$E(X_s|\mathcal{F}_t) \overset{\text{a. s.}}{=\!=} f(X_t, X_{t-1}, \cdots) \overset{\mathrm{d}}{=} f(X_{t+h}, X_{t+h-1}, \cdots) \overset{\text{a. s.}}{=\!=} E(X_{s+h}|\mathcal{F}_{t+h})$$

进一步，对固定的 $t_1<\cdots<t_k$，$k\in\mathbf{N}$，存在可测函数 f_1，…，f_k，使得

$$(E(X_{t_1+j}|\mathcal{F}_{t_1}), \cdots, E(X_{t_k+j}|\mathcal{F}_{t_k})) \overset{\text{a. s.}}{=\!=} (f_1(X_{t_1+j}, \cdots), \cdots, f_k(X_{t_k+j}, \cdots)).$$

上式右边依分布等于$(f_1(X_{t_1+j+h}, \cdots), \cdots, f_k(X_{t_k+j+h}, \cdots))$，而后者 a. s. 等于 $(E(X_{t_1+j+h}|\mathcal{F}_{t_1+h}), \cdots, E(X_{t_k+j+h}|\mathcal{F}_{t_k+h}))$.

现在可以证明$\{Z_t\}$是严平稳的了. 事实上，存在可测函数 f_1，…，f_k 使得

$$\begin{aligned}(Z_{t_1}, \cdots, Z_{t_k}) &= \sum_{j=0}^{\infty}(E(X_{t_1+j}|\mathcal{F}_{t_1}-1), \cdots, E(X_{t_k+j}|\mathcal{F}_{t_k}-1)) \\ &\overset{\text{a. s.}}{=\!=} \sum_{j=1}^{\infty}(f_1(X_{t_1+j-1}, \cdots), \cdots, f_k(X_{t_k+j-1}, \cdots)) \\ &\overset{\mathrm{d}}{=} \sum_{j=1}^{\infty}(f_1(X_{t_1+h+j-1}, \cdots), \cdots, f_k(X_{t_k+h+j-1}, \cdots)) \\ &\overset{\text{a. s.}}{=\!=} \sum_{j=0}^{\infty}(E(X_{t_1+h+j}|\mathcal{F}_{t_1+h-1}), \cdots, E(X_{t_k+h+j}|\mathcal{F}_{t_k+h-1})) \\ &= \sum (Z_{t_1+h}, \cdots, Z_{t_k+h}).\end{aligned}$$

这就证明了$\{Z_t\}$的严平稳性. 从而得到：

$$\begin{aligned}\mathrm{Var}\left(\frac{Z_1-Z_{T+1}}{\sqrt{T}}\right) &= \frac{1}{T}E(Z_1-Z_{T+1})^2 \\ &\leqslant \frac{2}{T}(EZ_1^2+EZ_{T+1}^2)\end{aligned}$$

$$=\frac{4EZ_1^2}{T},$$

当 $T\to\infty$ 时，上式收敛于 0.

现验证(ⅳ)，注意到

$$\begin{aligned}E(Y_t\mid\mathcal{F}_{t-1})&=\sum_{j=0}^{\infty}E(E(X_{t+j}\mid\mathcal{F}_t)-E(X_{t+j}\mid\mathcal{F}_{t-1})\mid\mathcal{F}_{t-1})\\&=\sum_{j=0}^{\infty}(E(X_{t+j}\mid\mathcal{F}_{t-1})-E(X_{t+j}\mid\mathcal{F}_{t-1}))=0(\text{a. s.})\end{aligned}$$

最后证明(ⅴ). 因为 Y_t 是严平稳的 L_2 鞅差序列，因此满足中心极限定理. 由分解

$$X_t=Y_t+Z_t-Z_{t+1}$$

272

可得

$$\frac{1}{\sqrt{T}}\sum_{t=1}^{T}X_t=\frac{1}{\sqrt{T}}\sum_{t=1}^{T}Y_t+R_T,$$

其中余项

$$R_T=\frac{Z_1-Z_{T+1}}{\sqrt{T}}$$

依 L_2 收敛于 0. 由 Slutaky 引理可得

$$\lim_{T\to\infty}P\Big(\frac{1}{\sqrt{T}}\sum_{t=1}^{T}X_t\leqslant x\Big)=\lim_{T\to\infty}P\Big(\frac{1}{\sqrt{T}}\sum_{t=1}^{T}Y_t\leqslant x\Big)=\Phi\Big(\frac{x}{\sigma}\Big),$$

对所有的 $x\in\mathbf{R}$ 和某个 $\sigma\geqslant 0$ 成立，并且

$$\lim_{n\to\infty}\operatorname{Var}\Big(\frac{1}{\sqrt{T}}\sum_{t=1}^{T}X_t\Big)=\lim_{n\to\infty}\operatorname{Var}\Big(\frac{1}{\sqrt{T}}\sum_{t=1}^{T}Y_t\Big).$$

因此，

$$\sigma^2=E(X_1^2)+2\sum_{h=0}^{\infty}E(X_1X_{1+h}),$$

在右边级数收敛时成立. 下面证明右边级数收敛，注意到

$$E\Big(\frac{1}{\sqrt{T}}\sum_{t=1}^{T}X_t\Big)^2=E(X_1^2)+2\sum_{h=1}^{T-1}\Big(1-\frac{h}{T}\Big)E(X_1X_{1+h}).$$

所以有

$$\begin{aligned}|E(X_1X_{1+h})|&=E(X_0E(X_h\mid\mathcal{F}_0))\\&\leqslant E(|X_0E(X_h\mid\mathcal{F}_0)|)\\&\leqslant\|X_0\|_2\|E(X_h\mid\mathcal{F}_0)\|_2\\&=\|X_0\|_2\|E(X_0\mid\mathcal{F}_{-h})\|_2.\end{aligned}$$

因此，

$$\sum_{h=1}^{\infty}|E(X_1X_{1+h})|\leqslant\|X_0\|_2\sum_{h=1}^{\infty}\|E(X_0\mid\mathcal{F}_{-h})\|_2<\infty.$$

■

8.2 金融时间序列回归模型

经典的多元线性回归模型刻画了 p 个解释变量(回归量)x_1，…，x_p 对依赖变量 Y 的影响，即假定

273

$$Y=\beta_0+\beta_1 x_1+\cdots+\beta_p x_p+\varepsilon,$$

其中 β_0，…，β_p 是未知的回归系数，ε 是零均值的误差项，在该模型中，假定回归量和系数都是非随机的. 为了估计上述模型，先收集 T 个观测向量$(Y_t, x_{t1}, \cdots, x_{tp})$，$t=1$，…，$T$. 对应的样本模型为

$$Y_t=\beta_0+\beta_1 x_{t1}+\cdots+\beta_p x_{tp}+\varepsilon_t,\quad t=1,\cdots,T.$$

取样得到的 T 个样本决定了随机变量 Y_t，x_{t1}，…，x_{tp}，$\varepsilon_t(t=1, \cdots, T)$之间的随机关系. 经典的回归模型假设误差项是独立同分布的，并服从正态分布. 但这种形式的回归模型不适用于金融数据集. 首先，需要金融模型中的回归量，比如收益，是随机的；其次，滞后值 $Y_{t-j}(j\in\mathbf{N})$通常作为解释变量出现；最后，误差项 ε_t 通常既不是独立的也不是正态的. 因此，需要考虑下列带随机回归量的一般回归模型，

$$Y_t=X_t'\beta+\varepsilon_t,\quad t=1,\cdots,T,\tag{8.2}$$

其中 X_t 是 d 维随机向量，$\beta\in\mathbf{R}^d$ 是固定的未知参数向量，ε_t 是随机误差项.

对回归量，本章作如下假定：

(R1) $\{X_t : t\in\mathbf{N}\}$是弱平稳的，即对所有的 t，有 $E(X_t)=E(X_1)$且 $E(X_tX_t')=E(X_1X_1')$，同时对所有 j 和 t，有 $E(X_{tj}^2)<\infty$.

接下来引入滤子$\mathcal{F}_t$，$t\geqslant 1$，它表示可能与回归量 X_t 有关的 t 时刻所收集到的信息. 设

(F1) X_{t-j}是$\mathcal{F}_t$ 可测的，$j\geqslant 0$.

(F2) Y_{t-j}是$\mathcal{F}_t$ 可测的，$j\geqslant 1$.

显然，可以取$\mathcal{F}_t=\sigma(X_{t-j}, Y_{t-1-j}: j\geqslant 0)$，但有时$\mathcal{F}_t$ 是给定的，这种情况下需要验证(F1)和(F2)是否成立.

关于误差项的假设如下：

(E1) $E(\varepsilon_t|\mathcal{F}_t)=0$(a. s.)，对所有的 $t\in\mathbf{N}$.

(E2) $E(\varepsilon_t^2|\mathcal{F}_t)=\sigma^2$(a. s.)，$t\in\mathbf{N}$，对某个常数 $\sigma^2\in(0, \infty)$.

在回归模型(8.2)中，ε_t 是预测值 $X_t'\beta$ 的随机扰动. (E1)表示对于给定的回归量，扰动项的条件均值为 0，即预测值是固定的. (E2)表明误差项的条件扩散是常量，与时间无关，在后面将讨论如何放宽这个条件.

下列结论简单却非常重要：ε_t 不是$\mathcal{F}_t$ 可测的，否则

$$\varepsilon_t^2=E(\varepsilon_t^2|F_t)=\sigma^2$$

对所有的 t 成立，则 ε_t^2 是一个常量时间序列，其边际方差 $\sigma^2=E(\varepsilon_t^2)=0$，这是一个矛盾. 从而得知 $Y_t=X_t'\beta+\varepsilon_t$ 不是$\mathcal{F}_t$ 可测的，否则 $\varepsilon_t=Y_t-X_t'\beta$ 是$\mathcal{F}_t$ 可测的. 易证所有的滞后值

274

ε_{t-j}，$j\geqslant 1$ 是$\mathcal{F}_t$ 可测的.

假定(E1)也表明：对不同的 t，ε_t' 是不相关的，事实上，对 $j\geqslant 1$，由 ε_{t-j}是$\mathcal{F}_t$ 可测的

且(E1)成立，从而有

$$E(\varepsilon_{t-j}\varepsilon_t)=E(E(\varepsilon_{t-j}\varepsilon_t|\mathcal{F}_t))$$
$$=E(\varepsilon_{t-j}E(\varepsilon_t|\mathcal{F}_t))=0,$$

但是，$\{\varepsilon_t\}$不是滤子$\{\mathcal{F}_t\}$的鞅差序列，因为是假定$E(\varepsilon_t|\mathcal{F}_t)=0$，而不是$E(\varepsilon_t|\mathcal{F}_{t-1})=0$. 但是可以通过扩大滤子使得$\{\varepsilon_t\}$变为一个鞅差序列，这样就可以运用强有力的极限定理来构建带随机回归量的回归模型的渐近假设检验和置信区间. 为此定义：

$$\mathcal{F}_t^*=\sigma(\mathcal{F}_t\cup\sigma(Y_t)),\quad t\in\mathbf{N}.$$

则下列性质成立：

(ⅰ) Y_t 是$\mathcal{F}_t^*$ 可测的.

(ⅱ) $\mathcal{F}_{t-1}^*=\sigma(\mathcal{F}_{t-1}\cup\sigma(Y_{t-1}))\subset\mathcal{F}_t$，因为$\mathcal{F}_{t-1}\subset\mathcal{F}_t$，$\sigma(Y_{t-1})\subset\mathcal{F}_t$.

(ⅲ) 由(ⅱ)可以计算条件期望，比如，由法则

$$E(Z|\mathcal{F}_{t-1}^*)=E(E(Z|\mathcal{F}_t)|\mathcal{F}_{t-1}^*),$$

可以计算$E(Z|F_{t-1}^*)$.

由性质(ⅲ)立刻得到

$$E(\varepsilon_t|\mathcal{F}_{t-1}^*)=E(E(\varepsilon_t|\mathcal{F}_t)|\mathcal{F}_{t-1}^*)=0,$$

这就验证了ε_t是一个$\mathcal{F}_t^*$ 鞅差序列.

▶**例 8.2.1** 著名的**资本资产定价模型(CAPM)**是一个线性回归模型. 假设一个经济活动(市场)有N个金融资产和一个银行账户，人们可以用相同的无风险利率r向银行存或借贷资金. 用R_i表示资产i在t时刻的对数收益，$i=1,\cdots,N$，$t=1,\cdots,T$. 假设市场存在一个投资组合，其对数收益为$R_t^{(M)}$. 记R为风险资产的对数收益，则差

$$R_e=R-r$$

称为**超额回报**. CAPM 模型的 Sharpe-Lintner 型是指超额回报$R_{ti}-r$满足下列线性回归模型

$$R_{ti}-r=\beta_0+\beta_i(R_t^{(M)}-r)+\varepsilon_{ti},\quad t=1,\cdots,T;i=1,\cdots,N.$$

回归系数β_i是资产的**贝塔因子(投资$\boldsymbol{\beta}$)**. 它们度量市场变化时资产回报的敏感度. 小于1的β的资产对市场波动不敏感，大于1的β的资产比投资组合有更大的风险. CAPM 理论表明$\beta_0=0$，扰动项是不相关的，且服从正态分布. 在实践应用时，通常需要用一个指标来表示投资组合. 且需要有一个确定的无风险利率，比如，国债的利率. ◀

275

8.2.1 最小二乘估计

估计回归模型最常用的方法是最小二乘法，即求使函数

$$Q_T(\beta)=\sum_{t=1}^{T}(Y_t-X_t'\beta)^2,\quad\beta\in\mathbf{R}^d$$

达到最小的参数值. 容易验证任一解$\hat{\beta}_T\in\mathbf{R}^d$ 满足下列正规方程

$$\boldsymbol{X}_T'\boldsymbol{X}_T\hat{\beta}_T=\boldsymbol{X}_T'\boldsymbol{Y},$$

其中$\boldsymbol{X}_T$是$T\times d$随机矩阵

$$\boldsymbol{X}_T = (X_1, \cdots, X_T)',$$

且

$$\boldsymbol{Y} = (Y_1, \cdots, Y_T)'.$$

注意到

$$\frac{1}{T}\boldsymbol{X}'_T\boldsymbol{X}_T = \frac{1}{T}\sum_{t=1}^{T} X_t X_t'$$

是 T 个秩 1 矩阵 $X_t X_t'(t=1, \cdots, T)$的均值. 如果 $\boldsymbol{X}'_T\boldsymbol{X}_T$ 是 a. s. 可逆的(至少对足够大的 T 是可逆的)，则可从正规方程中解得$\hat{\beta}_T$. 为此，加入下列假定以确保这一条件满足.

(R2)当 $T\to\infty$时，$\frac{1}{T}\boldsymbol{X}'_T\boldsymbol{X}_T\to\Sigma_{\boldsymbol{X}}$(依概率)，其中，$\Sigma_{\boldsymbol{X}}$ 是一个非随机的正定矩阵.

在某些实际问题中，甚至可以保证上述收敛是几乎处处收敛的，在这种情况下，可以找到某个集合 $A\subset\Omega$，使得当 $T\to\infty$时，$\frac{1}{T}\boldsymbol{X}_T(\omega)'\boldsymbol{X}_T(\omega)\to\Sigma_{\boldsymbol{X}}$，对所有的 $\omega\in A$ 均成立，剩下的 $\omega\in A^c$ 可能不成立. 下面可以假设上述条件是几乎处处收敛成立. 对足够大的 T，解得下列显示表达式

$$\hat{\boldsymbol{\beta}}_T = (\boldsymbol{X}'_T\boldsymbol{X}_T)^{-1}\boldsymbol{X}'_T\boldsymbol{Y}.$$

下面将推导出一个重要的公式. 由 $\boldsymbol{Y}=\boldsymbol{X}\boldsymbol{\beta}+\boldsymbol{\varepsilon}$ 可得

$$\begin{aligned}\hat{\boldsymbol{\beta}}_T &= (\boldsymbol{X}'_T\boldsymbol{X}_T)^{-1}\boldsymbol{X}'_T\boldsymbol{Y}\\ &= (\boldsymbol{X}'_T\boldsymbol{X}_T)^{-1}\boldsymbol{X}'_T(\boldsymbol{X}_T\boldsymbol{\beta}+\boldsymbol{\varepsilon})\\ &= \boldsymbol{\beta}+(\boldsymbol{X}'_T\boldsymbol{X}_T)^{-1}\boldsymbol{X}'_T\boldsymbol{\varepsilon},\end{aligned}$$

276

所以

$$\hat{\boldsymbol{\beta}}_T - \boldsymbol{\beta} = (\boldsymbol{X}'_T\boldsymbol{X}_T)^{-1}\sum_{t=1}^{T} X_t\varepsilon_t,$$

等价地有

$$\sqrt{T}(\hat{\boldsymbol{\beta}}_T - \boldsymbol{\beta}) = \left(\frac{1}{T}\boldsymbol{X}'_T\boldsymbol{X}_T\right)^{-1}\frac{1}{\sqrt{T}}\sum_{t=1}^{T} X_t\varepsilon_t.$$

假定(R2)确保了矩阵乘积的第一个因子$\left(\frac{1}{T}\boldsymbol{X}'_T\boldsymbol{X}_T\right)^{-1}$当 $T\to\infty$时依概率收敛于 $\Sigma_{\boldsymbol{X}}^{-1}$. 如果还知道当 $T\to\infty$时

$$\frac{1}{\sqrt{T}}\sum_{t=1}^{T} X_t\varepsilon_t \xrightarrow[T\to\infty]{d} N(\boldsymbol{0}, \boldsymbol{S}),$$

对某个矩阵 $\boldsymbol{S}$ 成立，则可以得到$\hat{\beta}_T$满足中心极限定理，因为此时有：当 $T\to\infty$时，

$$\sqrt{T}(\hat{\boldsymbol{\beta}}_T - \boldsymbol{\beta}) \xrightarrow[T\to\infty]{d} N\left(\boldsymbol{0}, \Sigma_X^{-1}\boldsymbol{S}\left(\Sigma_X^{-1}\right)'\right).$$

这意味着，对满足(R2)的回归量，必须寻找一个关于误差 ε_t 的条件，使得过程 $\xi_t = X_t\varepsilon_t$，$t\in\mathbf{N}$满足中心极限定理.

下列简单引理在接下来的内容中起着重要的作用.

引理 8.2.2 设假定(E1),(E2),(F1),(F2)以及(R1)成立,则过程 $\xi_t = X_t\varepsilon_t$, $t\in\mathbf{N}$ 是一个 $\mathcal{F}_t^*$ 鞅差序列,满足

$$\mathrm{Var}(\xi_t) = \sigma^2\Sigma_{\mathbf{X}}, \Sigma_{\mathbf{X}} = E(\boldsymbol{X}_t\boldsymbol{X}_t')$$

且

$$\mathrm{Cov}(\xi_s,\xi_t) = 0, \quad s < t.$$

证明 由 X_t, $\varepsilon_t\in L_2$ 知 $\xi_t = X_t\varepsilon_t\in L_1$. 又由 $E(\varepsilon_t|\mathcal{F}_t)=0$ 得到

$$\begin{aligned}E(X_t\varepsilon_t|\mathcal{F}_{t-1}^*) &= E(E(X_t\varepsilon_t|\mathcal{F}_t)|\mathcal{F}_{t-1}^*)\\ &= E(X_tE(\varepsilon_t|\mathcal{F}_t)|\mathcal{F}_{t-1}^*) = 0,\end{aligned}$$

且

$$\begin{aligned}\mathrm{Var}(X_t\varepsilon_t) &= E(\varepsilon_t^2X_tX_t')\\ &= E(E(\varepsilon_t^2X_tX_t'|\mathcal{F}_t))\\ &= E(X_tX_t'E(\varepsilon_t^2|\mathcal{F}_t))\\ &= \sigma^2E(X_tX_t') = \sigma^2\Sigma_{\boldsymbol{X}},\end{aligned}$$

最后一个等式用到假定(R1). ■ 277

8.3 鞅差阵列的极限定理

下面列举一些鞅差阵列的基本性质,这些结果推广了鞅差序列的结果.

定义 8.3.1 一个随机变量构成的矩阵列 X_{nk}: $(\Omega, \mathcal{F}, P)\to(\mathbf{R}, \mathcal{B})$ $(1\leqslant k\leqslant r_n$, $n\geqslant 1)$ 称为关于单调非减 σ 代数流 $\{\mathcal{F}_{nk}$: $1\leqslant k\leqslant r_n$, $n\geqslant 1\}$ 的一个鞅差阵列,若对于每个 $n\in\mathbf{N}$,序列

$$X_{n1},\cdots,X_{nrn}$$

是关于滤子

$$\mathcal{F}_{n1}\subset\cdots\subset\mathcal{F}_{nr}$$

的一个鞅差序列. 即

(i) $\mathcal{F}_{n0}=\{\emptyset, \Omega\}$, $\mathcal{F}_{nk}\subset\mathcal{F}_{n,k+1}\subset\cdots$, $1\leqslant k\leqslant r_n$, $n\geqslant 1$.

(ii) X_{nk} 是 $\mathcal{F}_{nk}$ 可测的且 $E|X_{nk}|<\infty$, $1\leqslant k\leqslant r_n$, $n\geqslant 1$.

(iii) $E(X_{nk}|\mathcal{F}_{n,k-1})=0$, 对 $1\leqslant k\leqslant r_n$, $n\geqslant 1$.

鞅差阵列在一个相当一般的条件下满足下列弱大数定律.

定理 8.3.2 设 $\{X_{Tt}\}=\{X_{Tt}$: $1\leqslant t\leqslant T$, $T\geqslant 1\}$ 是一个 $\mathcal{F}_{T,t}$ 鞅差阵列,且下列两个条件之一满足:

(i) 存在某个 $\sigma^2\in[0, \infty)$ 使得当 $T\to\infty$ 时有

$$\frac{1}{T}\sum_{t=1}^{T}\mathrm{Var}(X_{Tt})\to\sigma^2;$$

(ii) 存在某个 $\sigma^2\in[0, \infty)$ 使得当 $T\to\infty$ 时有

$$\frac{1}{T}\sum_{t=1}^{T}\sigma_{Tt}^2\to\sigma^2, 在 L_1 中,$$

其中 $T\to\infty$，$\sigma_{Tt}^2=E(X_{Tt}^2\mid\mathcal{F}_{T,t-1})$. 则下列弱大数定律成立：当 $T\to\infty$时有

278

$$\frac{1}{T}\sum_{t=1}^{T}X_{Tt}\xrightarrow{P}0.$$

证明 由切比雪夫不等式知

$$P\left(\left|\frac{1}{T}\sum_{t=1}^{T}X_{Tt}\right|>\varepsilon\right)\leqslant\frac{E\left(\frac{I}{T}\sum_{t=1}^{T}X_{Tt}\right)^2}{\varepsilon^2}$$

$$\leqslant\frac{\frac{I}{T^2}\sum_{t=1}^{T}E(X_{Tt}^2)}{\varepsilon^2},$$

其中用到了鞅差序列是以原点为中心，且不相关的事实. 如果(ⅱ)满足，由于 $T^{-1}\sum_{t=1}^{T}\sigma_{Tt}^2\to\sigma^2$（依 L_1 收敛）意味着当 $T\to\infty$时有 $a_T=E\left(T^{-1}\sum_{t=1}^{T}\sigma_{Tt}^2\right)\to\sigma^2$，所以$\{a_T\}$是有界的(设界为常数 c)，所以有

$$\frac{1}{T^2}\sum_{t=1}^{T}E(X_{Tt}^2)=\frac{1}{T}E\left[\frac{1}{T}\sum_{t=1}^{T}E(X_{Tt}^2\mid\mathcal{F}_{T,t-1})\right]$$

$$=\frac{1}{T}E\left(\frac{1}{T}\sum_{t=1}^{T}\sigma_{Tt}^2\right)\leqslant\frac{c}{T}.$$

(ⅰ)是充分条件的证明留给读者. ■

下面通过估计对数收益的均值和波动率的例子来说明定理 8.3.2.

▶**例 8.3.3** 假设 R_1，R_2，…是一项资产的对数收益. 为了估计以 R_1，…，R_T 为观测值的总体的均值和波动率，考虑下列基本模型

$$R_t=\mu+\varepsilon_t,\quad t\in\mathbf{N},$$

其中 $\mu\in\mathbf{R}$ 是一个常量，$\{\varepsilon_t\}$是关于自然滤子$\mathcal{F}_t=\sigma(\varepsilon_1,\cdots,\varepsilon_t)$的一个鞅差矩阵列，满足

$$E(\varepsilon_t^2\mid\mathcal{F}_{t-1})=\sigma^2,\quad E(\varepsilon_t^4\mid\mathcal{F}_{t-1})=\gamma_4<\infty,\text{a. s.}$$

对所有的 t 成立，其中 γ_4 和 σ^2 是常数. 显然 $E(R_t\mid\mathcal{F}_{t-1})=\mu+E(\varepsilon_t\mid\mathcal{F}_{t-1})=\mu$，因此 $E(R_t)=\mu$. 由此可得

$$\hat{\mu}_T=\frac{1}{T}\sum_{t=1}^{T}R_t$$

是 μ 的无偏估计量. 由 $\hat{\mu}_t-\mu=T^{-1}\sum_{t=1}^{T}\varepsilon_t$，根据定理 8.3.2 可得当 $T\to\infty$时

$$\hat{\mu}_T-\mu\xrightarrow{P}0\Leftrightarrow\hat{\mu}_T\xrightarrow{P}\mu,$$

从而$\hat{\mu}_t$是 μ 的弱一致估计量.

为了估计 σ^2，首先考虑 μ 是已知的情形. 此时由于$(R_t-\mu)^2=\varepsilon_t^2$ 所以 $E(R_t-\mu)^2=\sigma^2$，则自然想到用

$$\widetilde{\sigma}_T^2 = \frac{1}{T}\sum_{t=1}^{T}(R_t - \mu)^2$$

279

来估计 σ^2. 下面证明$\{\varepsilon_t^2 - \sigma^2: t \geqslant 1\}$是一个鞅差序列. 事实上，由假设知

$$E|\varepsilon_t^2 - \sigma^2| \leqslant E\varepsilon_t^2 + \sigma^2 < \infty$$

且

$$E(\varepsilon_t^2 - \sigma^2 | \mathcal{F}_{t-1}) = E(\varepsilon_t^2 | \mathcal{F}_{t-1}) - \sigma^2 = 0(\text{a. s.})$$

为了能运用定理 8.3.2，注意到

$$\begin{aligned}\frac{1}{T}\sum_{t=1}^{T}E[(\varepsilon_t^2 - \sigma^2)^2 | \mathcal{F}_{t-1}] &= \frac{1}{T}\sum_{t=1}^{T}E(\varepsilon_t^4 - 2\varepsilon_t^2\sigma^2 + \sigma^4 | \mathcal{F}_{t-1}) \\ &= \gamma_4 - \sigma^4 < \infty.\end{aligned}$$

由此可得当 $T \to \infty$ 时有$\widetilde{\sigma}_T^2 - \sigma^2 \xrightarrow{P} 0$.

若 μ 是未知的，用$\hat{\mu}_t$代替 μ，考虑

$$\hat{\sigma}_T^2 = \frac{1}{T}\sum_{t=1}^{T}(R_t - \hat{\mu}_T)^2.$$

$\hat{\sigma}_T^2$ 的一致性根据下列事实得到：

$$\begin{aligned}\hat{\sigma}_T^2 &= \frac{1}{T}\sum_{t=1}^{T}(R_t - \mu + \mu - \hat{\mu}_T)^2 \\ &= \frac{1}{T}\sum_{t=1}^{T}(R_t - \mu)^2 + 2(\mu - \hat{\mu}_T)\frac{1}{T}\sum_{t=1}^{T}(R_t - \mu) + (\mu - \hat{\mu}_T)^2.\end{aligned}$$

第一项是$\widetilde{\sigma}_T^2$. 由$\hat{\mu}_T - \mu \xrightarrow{P} 0$ 和$\frac{1}{T}\sum_{t=1}^{T}(R_t - \mu) \xrightarrow{P} 0$ 可知第二项依概率收敛于 0. 最后一项，$(\mu - \hat{\mu}_T)^2$也依概率收敛于 0. 由 Slutzky 引理知：当 $T \to \infty$ 时，$\hat{\sigma}_T^2 \xrightarrow{P} \sigma^2$. ◀

下面继续研究广义回归模型. 下面的命题表明，如果每个观测值乘以过去值的函数，那么鞅差阵列的性质将会保留.

推论 8.3.4 设$\{X_t: t \geqslant 0\}$是一个$\mathcal{F}_t$ 鞅差阵列，满足 $X_t \in L_p$ 对所有的 t 成立，令 g 是定义在 $\mathbf{R}^\infty$ 上的函数，满足

$$Y_{t-1} = g(X_{t-1}, X_{t-2}, \cdots) \in L_q, \text{对所有的 } t,$$

其中$\frac{1}{p} + \frac{1}{q} = 1$，$p$，$q \in [1, \infty)$. 则$\{X_t Y_{t-1}: t \geqslant 1\}$是一个鞅差阵列. 进一步，$\{X_t: t \geqslant 1\}$和$\{Y_{t-1}: t \geqslant 1\}$是不相关的.

证明 由 Hölder 不等式知

$$E|X_t Y_{t-1}| \leqslant \|X_t\|_p \|Y_{t-1}\|_q < \infty.$$

280

进一步，$E(X_t Y_{t-1} | \mathcal{F}_{t-1}) = Y_{t-1}E(X_t | \mathcal{F}_{t-1}) = 0$，a. s.，且

$$\text{Cov}(X_t, Y_{t-1}) = E(X_t Y_{t-1}) = E(Y_{t-1}E(X_t | \mathcal{F}_{t-1})) = 0.$$ ■

综合定理 8.3.2 和命题 8.3.4 可得到下列关于鞅变换的一个有趣结论.

定理 8.3.5 设$\{X_t\}$是一个$\mathcal{F}_t$ 鞅，φ_t 是$\mathcal{F}_t$ 可预报的，如果当 $T \to \infty$ 时有

$$\frac{1}{T}\sum_{t=1}^{T}\varphi_t^2 E[(\Delta X_t)^2 \mid \mathcal{F}_{t-1}] \xrightarrow{P} \sigma^2,$$

其中 $\sigma^2 \in (0, \infty)$，那么

$$\frac{1}{T}\int_0^T \varphi_t \mathrm{d}X_t \underset{T\to\infty}{\xrightarrow{P}} 0.$$

证明 注意到

$$\frac{1}{T}\int_0^T \varphi_t \mathrm{d}X_t = \frac{1}{T}\sum_{t=1}^{T}\varphi_t\xi_t = \frac{1}{T}\sum_{t=1}^{T}\widetilde{\varphi}_{t-1}\xi_t,$$

其中 $\xi_t = \Delta X_t = X_t - X_{t-1}$ 是一个 $\mathcal{F}_t$ 鞅差序列，且 $\widetilde{\varphi}_{t-1} = \varphi_t$ 是 $\mathcal{F}_{t-1}$ 可测的. 因此，$\widetilde{\varphi}_{t-1}\xi_t$ 是鞅差序列，由定理 8.3.2 知定理 8.3.5 的结论成立. ■

在相当一般的条件下，鞅差阵列满足下列中心极限定理，详细的证明可参考相关文献.

定理 8.3.6 设 $\{X_{Tt}\} = \{X_{Tt}: 1 \leqslant t \leqslant T, T \geqslant 1\}$ 是一个鞅差阵列，满足 $\sigma_{Tt}^2 = \mathrm{Var}(X_{Tt}) < \infty$ 对所有的 t，T 成立，且

$$\mathrm{Var}\Big(\sum_{t=1}^{T} X_{Tt}\Big) = \sum_{t=1}^{T}\sigma_{Tt}^2 = 1.$$

如果

(i) 当 $T \to \infty$ 时，$\sum_{t=1}^{T} X_{Tt}^2 \xrightarrow{P} 1$

且

(ii) 当 $T \to \infty$ 时，$\max_{1\leqslant t\leqslant T} |X_{T_t}| \xrightarrow{P} 0$

成立，则当 $T \to \infty$ 时，有

281

$$\sum_{t=1}^{T} X_{Tt} \xrightarrow{d} N(0,1).$$

注 8.3.7 关于鞅差阵列的中心极限定理有多种形式. 定理 8.3.6 的充分条件使人们能够用相对初等的方法确定金融数据统计量的渐近正态性. 特别地，可以用下列 Linerberg 条件来代替(ii)：

$$\sum_{i=1}^{T} E(\xi_{Ti}^2 \mathbf{1}(|\xi_{Tj}| > \varepsilon) \mid \mathcal{F}_{T,i-1}) \to 0, \tag{8.3}$$

对任何 $\varepsilon > 0$ 成立. 附录 B 中列出了金融统计中常用的几种鞅差阵列的中心极限定理.

对于严平稳的鞅差序列，有下列简化结果.

定理 8.3.8 设 $\{X_t: t \in \mathbf{N}\}$ 是一个严平稳的 L_2 鞅差序列，即

$$\sigma^2 = E(X_t^2) < \infty$$

且

$$E(X_t \mid \mathcal{F}_{t-1}) = 0,$$

对所有的 t 成立，其中对 $t \in \mathbf{N}$，$\mathcal{F}_t = \sigma(X_s: s \leqslant t)$，且 $F_0 = \{\varnothing, \Omega\}$. 那么，当 $T \to \infty$ 时

$$\frac{1}{\sqrt{T}}\sum_{t=1}^{T} X_t \xrightarrow{d} N(0, \sigma^2).$$

证明 定义随机变量

$$\xi_t = E(X_t^2 | \mathcal{F}_{t-1}), t \geqslant 1,$$

注意到 $\xi_t = f(X_{t-1}, X_{t-2}, \cdots)$，其中 f 是定义在 $\mathbf{R}^{\mathbf{N}_0}$ 上的 Borel 可测函数，$\{\xi_t\}$ 是一个严平稳过程，因此满足遍历定理．则有，当 $T \to \infty$ 时

$$\frac{1}{T}\sum_{t=1}^{T} E(X_t^2 | \mathcal{F}_{t-1}) = \frac{1}{T}\sum_{t=1}^{T} \xi_t \xrightarrow{P} \sigma^2,$$

进一步，由严平稳性有：

$$\frac{1}{T}\sum_{t=1}^{T} E(X_t^2 \mathbf{1}_{\{|X_t| > \varepsilon\sqrt{T}\}}) = E(X_1^2 \mathbf{1}_{\{|X_1| > \varepsilon\sqrt{T}\}}).$$

由控制收敛定理知，当 $T \to \infty$ 时，上式收敛于 0，这就验证了条件(8.3)式. ■ 282

为了说明如何应用定理 8.3.6，下面将会看到，当收益构成一个鞅差序列时，渐近性使得估计(对数)收益的均值的统计检验和置信区间仍然是有效的.

▶**例** 8.3.9 在本章的开始，我们回顾了均值为 μ，波动率为 σ 的独立同分布的收益 $R_1, \cdots, R_T$ 的中心极限定理. 下面将会看到，在例 8.3.3 的条件下，

$$\sqrt{T}(\hat{\mu}_T - \mu) = \frac{1}{\sqrt{T}}\sum_{t=1}^{T}(R_t - \mu) \xrightarrow[T\to\infty]{d} N(0, \sigma^2) \tag{8.4}$$

仍然成立. 事实上，令

$$X_{Tt} = \frac{R_t - \mu}{\sqrt{T} \cdot \sigma}, 1 \leqslant t \leqslant T, T \in \mathbf{N},$$

则 $\dfrac{\sqrt{T}(\hat{\mu}_T - \mu)}{\sigma} = \sum_{t=1}^{T} X_{Tt}$. 显然，$\mathrm{Var}\left(\sum_{t=1}^{T} X_{Tt}\right) = 1$. 下面验证定理 8.3.6 的条件(ⅱ)，根据例 8.3.3，当 $T \to \infty$ 时，

$$\sum_{t=1}^{T} X_{Tt}^2 = \frac{1}{T\sigma^2}\sum_{t=1}^{T} \varepsilon_t^2 \to 1,$$

因此，由一致有界性和 Markov 不等式可得条件(ⅱ)成立. 从而当 $T \to \infty$ 时，有

$$\begin{aligned} P(\max_{1\leqslant t\leqslant T} \|R_t - \mu\| > \varepsilon\sqrt{T}\sigma) &= P\left(\bigcup_{t=1}^{T}\{|\varepsilon_t| > \varepsilon\sqrt{T}\sigma\}\right) \\ &\leqslant \sum_{t=1}^{T} P\left(|\varepsilon_t| > \varepsilon\sqrt{T}\sigma\right) \\ &\leqslant \sum_{t=1}^{T} \frac{E|\varepsilon_t|^4}{\varepsilon^4 T^2 \sigma^4} \\ &= \frac{E(\varepsilon_1^4)}{T\varepsilon^4\sigma^4} \to 0, \end{aligned}$$

(8.4)式得证. ◀

8.4 渐近性

在开始本节内容之前，先回过头来进一步观察假定(R2)，它为

$$\frac{1}{T}\sum_{t=1}^{T}X_tX_t' \xrightarrow{P} \Sigma_X, \quad T\to\infty.$$

283

首先，随机矩阵序列$\{A, A_T\}$依概率收敛是指当$T\to\infty$时，$P(\|A_T-A\|>\varepsilon)\to 0$，其中$\|.\|$表示矩阵范数，它等价于每个元素依概率收敛. 因此，不妨假设维数$d=1$，此时条件变为：当$T\to\infty$时

$$\frac{1}{T}\sum_{t=1}^{T}X_t^2 \xrightarrow{P} \sigma^2 = E(X_1^2) \in (0,\infty), \tag{8.5}$$

对独立同分布的回归量，该条件显然成立. 在假定(E2)下，$\{X_t^2-\sigma^2: t\geqslant 1\}$是一个$\mathcal{F}_t$鞅差序列，由例 8.3.3 知(8.5) 式也成立.

一个更重要的情形是带鞅差序列误差项的一阶自回归模型，即

$$X_t = \rho X_{t-1} + \varepsilon_t, \quad t\geqslant 1, \tag{8.6}$$

其中$\{\varepsilon_t: t\in\mathbf{Z}\}$是一个$\mathcal{F}_t$鞅差序列，且$|\rho|<1$. 在例 3.4.10 中已经看到上述方程有一个平稳解，即

$$X_t = \sum_{i=0}^{\infty}\rho^i\varepsilon_{t-i}, \quad t\in\mathbf{Z}.$$

因此，如果取随机初始值$X_0 = \sum_{i=0}^{\infty}\rho^i\varepsilon_{-i}$，并令

$$X_t = \sum_{i=0}^{\infty}\rho^i\varepsilon_{t-1}, \quad t\geqslant 1,$$

则可以得到一个解$\{X_t: t\in\mathbf{N}_0\}$满足弱平稳条件

$$E(X_t^i) = E(X_1^i) \text{ 对所有的 } t \text{ 以及 } i=1,2 \text{ 成立}. \tag{8.7}$$

命题 8.4.1 设模型(8.6)成立，如果对所有的t，$E(\varepsilon_t^2|\mathcal{F}_{t-1})=\sigma^2>0$，$E(\varepsilon_t^4|\mathcal{F}_{t-1})=\gamma^4$，其中$\sigma^2$，$\gamma^4<\infty$是常数，则当$T\to\infty$时，

$$\frac{1}{T}\sum_{t=1}^{T}X_t^2 \xrightarrow{P} \frac{\sigma^2}{1-\rho^2}.$$

证明 由

284

$$X_t^2 = \rho^2X_{t-1}^2 + 2\rho X_{t-1}\varepsilon_t + \varepsilon_t^2,$$

知

$$\begin{aligned}\frac{1}{T}\sum_{t=1}^{T}X_t^2 &= \frac{\rho^2}{T}\sum_{t=1}^{T}X_{t-1}^2 + \frac{2\rho}{T}\sum_{t=1}^{T}X_{t-1}\varepsilon_t + \frac{1}{T}\sum_{t=1}^{T}\varepsilon_t^2 \\ &=\rho^2\left(\frac{1}{T}\sum_{t=1}^{T}X_t^2\right) + \rho^2\left(\frac{X_0^2}{T}-\frac{X_T^2}{T}\right) + \frac{2\rho}{T}\sum_{t=1}^{T}X_{t-1}\varepsilon_t + \frac{1}{T}\sum_{t=1}^{T}\varepsilon_t^2 \\ &=\rho^2\left(\frac{1}{T}\sum_{t=1}^{T}X_t^2\right) + A_T + B_T + C_T.\end{aligned}$$

所以

$$(1-\rho^2)\frac{1}{T}\sum_{t=1}^{T}X_t^2 = A_T + B_T + C_T.$$

显然，对 $i=0$，…，T，当 $T\to\infty$ 时有

$$P\left(\rho^2 \left|\frac{X_i^2}{T}\right| > \varepsilon\right) \leqslant \rho^2 \frac{EX_0^2}{T\varepsilon} \to 0,$$

由此可得当 $T\to\infty$ 时 $A_T \xrightarrow{P} 0$. 进一步，由命题 8.3.4 知，$\varepsilon_t X_{t-1}$ 是一个 $\mathcal{F}_t$ 鞅差序列. 根据(8.7)式知，其边际方差为

$$\begin{aligned}\operatorname{Var}(\varepsilon_t X_{t-1}) &= E(\varepsilon_t^2 X_{t-1}^2)\\ &= E(X_{t-1}^2 E(\varepsilon_t^2 \mid \mathcal{F}_{t-1}))\\ &= \sigma^2 E(X_{t-1}^2)\\ &= \sigma^2 E(X_1^2) > 0,\end{aligned}$$

因此，定理 8.3.2 的条件(ⅰ)满足，从而当 $T\to\infty$ 时，$B_T \xrightarrow{P} 0$ 成立. 最后，考虑 C_T. 随机变量 $\varepsilon_t^2-\sigma^2(t\geqslant 1)$ 是鞅差序列，其方差为

$$\operatorname{Var}(\varepsilon_t^2-\sigma^2) = E(\varepsilon_t^2-\sigma^2)^2 = \gamma_4-\sigma^4 \in [0,\infty).$$

再一次应用定理 8.3.2 可得当 $T\to\infty$ 时，$C_T \xrightarrow{P} \sigma^2$.

为了建立最小二乘估计量 $\hat{\beta}_t$ 的中心极限定理，还需要在假定(R1)和(R2)的基础上增加条件. ■

定理 8.4.2 设(R1)、(R2)、(F1)、(F2)、(E1)、(E2)成立，且下列条件成立：

(SM)存在一个 $\delta>0$ 使得 $E|X_t\varepsilon_t|^{2+\delta}=E|X_1\varepsilon_1|^{2+\delta}<\infty$，对所有 t 成立，且 $E(X_1^4\varepsilon_1^2)=E(X_t^4\varepsilon_t^2)$ 对所有 t 也成立.

285

那么最小二乘估计量 $\hat{\beta}_t$ 是渐近正态的，即当 $T\to\infty$ 时，

$$\sqrt{T}(\hat{\boldsymbol{\beta}}_T-\boldsymbol{\beta}) \xrightarrow[T\to\infty]{d} N\left(0,\sigma^2\Sigma_X\right).$$

证明 当 $d=1$ 时，有

$$\sqrt{T}(\hat{\boldsymbol{\beta}}_T-\boldsymbol{\beta}) = \left(\frac{1}{T}\sum_{t=1}^{T} X_t^2\right)^{-1} \frac{1}{\sqrt{T}}\sum_{t=1}^{T} X_t\varepsilon_t.$$

将定理 8.3.6 应用于随机变量矩阵列

$$X_{Tt} = \frac{X_t\varepsilon_t}{\sqrt{T}\sigma\ \sqrt{\Sigma_x}}, 1\leqslant t\leqslant T, T\geqslant 1,$$

因为 $d=1$，所以 $\Sigma_x=\Sigma_X$. 在引理 8.2.2 中已经知道 $X_t\varepsilon_t$ 是一个 $\mathcal{F}_t^*$ 鞅差序列. 所以易证 X_{Tt} 是一个 $\mathcal{F}_t^*$ 鞅差阵列. 特别地，$\operatorname{Var}\left(\sum_{t=1}^{T} X_{Tt}\right)=1$ 对所有的 T 成立. 下面证明

(ⅰ) $\frac{1}{T}\sum_{t=1}^{T} X_t^2\varepsilon_t^2 \xrightarrow{P} \sigma^2\Sigma_X, T\to\infty$,

(ⅱ) $P(\max_{1\leqslant t\leqslant T}|X_t\varepsilon_t|>\varepsilon\eta\sqrt{T})\to 0$，$T\to\infty$，其中 $\eta=\sigma\sqrt{\Sigma_x}$.

首先验证(ⅱ). 与例 8.3.9 类似，根据一致有界性和 Markov 不等式，此时的函数是 $x\to x^{2+\delta}$，知当 $T\to\infty$ 时有

$$P\Big(\max_{1\leqslant t\leqslant T}|X_t\varepsilon_t|>\varepsilon\eta\sqrt{T}\Big)=P\Big(\bigcup_{t=1}^{T}\big\{|X_t\varepsilon_t|>\varepsilon\eta\sqrt{T}\big\}\Big)$$
$$\leqslant\sum_{t=1}^{T}P(|X_t\varepsilon_t|>\varepsilon\eta\sqrt{T})$$
$$\leqslant\sum_{t=1}^{T}\frac{E|X_t\varepsilon_t|^{2+\delta}}{(\varepsilon\eta)^{2+\delta}T^{1+\delta/2}}$$
$$=\frac{E|X_1\varepsilon_1|^{2+\delta}}{(\varepsilon\eta)^{2+\delta}T^{\delta/2}}\to 0,$$

这就证明了(ⅱ).(ⅰ)的验证稍显复杂.注意到

$$X_t^2\varepsilon_t^2-\sigma^2\Sigma_X=X_t^2(\varepsilon_t^2-\sigma^2)+\sigma^2(X_t^2-\Sigma_X),$$

由此可得分解

$$\frac{1}{T}\sum_{t=1}^{T}\Big(X_t^2\varepsilon_t^2-\sigma^2\Sigma_X\Big)=\frac{1}{T}\sum_{t=1}^{T}X_t^2(\varepsilon_t^2-\sigma^2)+\sigma^2\frac{1}{T}\sum_{t=1}^{T}\Big(X_t^2-\Sigma_X\Big).$$

286

由假设(R2)知,右边第二项当 $T\to\infty$ 时,依概率收敛于 0.将鞅差序列的大数定律应用于第一项.由(R1)和(SM)有

$$E|X_t^2(\varepsilon_t^2-\sigma^2)|\leqslant E|X_t^2\varepsilon_t^2|+E|X_t^2|\sigma^2$$
$$\leqslant\sqrt{EX_t^4}\sqrt{E\varepsilon_t^4}+E|X_t^2|\sigma^2<\infty.$$

进一步,由假设(E2),可得 $X_t^2(\varepsilon_t^2-\sigma^2)\geqslant 1$ 是$\mathcal{F}_t$ 鞅差序列,所以

$$E(X_t^2(\varepsilon_t^2-\sigma^2)\,|\,\mathcal{F}_{t-1}^*)=E(X_t^2\cdot E(\varepsilon_t^2-\sigma^2\,|\,\mathcal{F}_{t-1})\,|\,\mathcal{F}_{t-1}^*)=0.$$

现在验证

$$\frac{1}{T}\sum_{t=1}^{T}X_t^4(\varepsilon_t^2-\sigma^2)^2\xrightarrow[T\to\infty]{}c<\infty,\text{在 } L_1 \text{ 中},$$

其中 c 是常数.由(SM)知

$$E\Big(\frac{1}{T}\sum_{t=1}^{T}E(X_t^4(\varepsilon_t^2-\sigma^2)^2\,|\,\mathcal{F}_{t-1}^*)\Big)=E(X_1^4(\varepsilon_1^2-\sigma^2)^2).\tag{8.8}$$

因为 $E(X_1^4\varepsilon_1^2)=E(X_1^4E(\varepsilon_1^2\,|\,\mathcal{F}_1))=\sigma^2E(X_1^4)$,所以

$$E(X_1^4(\varepsilon_1^2-\sigma^2)^2)=E(X_1^4\varepsilon_1^4-2X_1^4\varepsilon_1^2\sigma^2+\sigma^4X_1^4)$$
$$=\sigma^2E(X_1^4\varepsilon_1^4)-2\sigma^4E(X_1^4)+\sigma^4E(X_1^4).$$

其中用到了 X_1 的$\mathcal{F}_1$ 可测性(见条件(F1)).所以(8.8)式的左边收敛于 $c=E(X_1^4\varepsilon_1^4)-\sigma^4E(X_1^4)<\infty$.所以(ⅰ)得证. ■

8.5 密度估计和非参数回归

第 1 章已经讨论了单变量样本的 Rosenblatt-Parzen 核估计.但是在金融应用中,经常涉及的是多变量的情形,比如分析一个投资组合的对数收益.不论是在实际应用中,还是在分析结果的推导中,非参数估计的偏差始终是一个问题.为了简化讨论,在理论分析时常常只讨论独立同分布的情形.然而,独立同分布得到的结果是可以推广到有相关性的时

间序列上去的，第 9 章将会讨论这一问题，在那里将讨论局部多项式估计. 287

8.5.1 多变量密度估计

第 1 章已经讨论了单变量非参数密度估计，现推广到多变量的情形，这既是实证的需要，也有助于更好地理解非参数回归. 设

$$X_t = (X_{t1}, \cdots, X_{td}), \quad t = 1, \ldots, T$$

是独立同分布的 d 维随机向量，$d \in \mathbf{N}$，其密度函数为 f. 设 L：$\mathbf{R} \to [0, \infty)$为单变量核，满足

$$\int L(z)\mathrm{d}z = 1, \quad \int zL(z)\mathrm{d}z = 0, \tag{8.9}$$

$$L_2 = \int L(z)^2 \mathrm{d}z < \infty, \quad \widetilde{L}_2 = \int z^2 L(z)\mathrm{d}z < \infty. \tag{8.10}$$

又设 h_1，…，$h_d > 0$ 为带宽，使用它们得到缩放光滑核 $h_j^{-1} L(x/h_j)$，$x \in \mathbf{R}$，$j = 1$，…，d. 值得一提的是，对每个 j，也可以使用一个不同的光滑核 L_j，除了多一些下标外，推导过程完全一样. 因此，为了简化记号，规定 $L_j = L$，$j = 1$，…，d.

现将 d 个缩放单变量核组合成缩放乘积核

$$\frac{1}{h_1 \ldots h_d} K\left(\frac{x}{h}\right) = \frac{1}{h_1 \ldots h_d} \prod_{j=1}^{d} L\left(\frac{x_j}{h_j}\right), \quad x = (x_1, \cdots, x_d) \in \mathbf{R}^d. \tag{8.11}$$

乍看之下左边的记号有点问题，但如果令 $h = (h_1, \ldots, h_d)$，对向量 $x \in \mathbf{R}^d$ 定义 x/h 为

$$(x_1/h_1, \cdots, x_d/h_d),$$

且定义 K：$\mathbf{R}^d \to \mathbf{R}$ 为

$$K(x_1, \cdots, x_d) = \prod_{j=1}^{d} L(x_j), \quad (x_1, \cdots, x_d) \in \mathbf{R}^d.$$

则(8.1)式左边是有定义的. 注意，乘积核定义了一个 d 维分布，若 $Z_j \sim L$，$j = 1$，…，d，则其独立坐标的边际分布为 $h_j Z_j$.

f 的多元核密度估计定义为

$$\hat{f}_T(x) = \frac{1}{T h_1 \cdots h_d} \sum_{t=1}^{T} K\left(\frac{x - X_t}{h}\right), \quad x \in \mathbf{R}^d.$$

这里对每个维度使用了相同的带宽 $h_j = h(j = 1$，…，$d)$时，公式简化了. 但是在实际中 d 个变量通常有不同的缩放，这就需要分别选择带宽来考虑. 288

对于单变量的情形，研究其估计量的偏差、方差以及均方误差(MSE)是有趣的.

命题 8.5.1 设 f 的三阶偏导数存在且所有偏导数一致有界，则

$$\mathrm{Bias}(\hat{f}_T(x); f(x)) = \frac{\widetilde{L}_2}{2} \sum_{j=1}^{d} h_j^2 \frac{\partial^2 f}{\partial x_j^2} + O\Big(\sum_{j=1}^{d} h_j^3\Big) \tag{8.12}$$

$$\mathrm{Var}(\hat{f}_T(x)) = \frac{1}{T h_1 \cdots h_d} \Big\{ L_2^d f(x) + O\Big(\sum_{j=1}^{d} h_j^2\Big) \Big\} \tag{8.13}$$

且

$$\mathrm{MSE}(\hat{f}_T(x);f(x)) = O\Big(\Big(\sum_{j=1}^d h_j^2\Big)^2 + \frac{1}{Th_1\cdots h_d}\Big).$$

证明 首先计算核密度估计量的偏差. 通过 d 个代换 $u_j=(z_j-x_j)/h_j$，$j=1$，…，d，得到

$$\begin{aligned}E(\hat{f}_T(x))-f(x) &= E\Big(\frac{1}{h_1\cdots h_d}K\Big(\frac{X_1-x}{h}\Big)-f(x)\Big)\\ &= \frac{1}{h_1\cdots h_d}\int K\Big(\frac{z-x}{h}\Big)f(z)\mathrm{d}z - f(x)\\ &= \int K(u)[f(x+hu)-f(x)]\mathrm{d}u,\end{aligned}$$

其中 $\mathrm{d}u=\mathrm{d}u_1\cdots\mathrm{d}u_d$. 由泰勒展开至三阶项知上述最后一个表达式可以写成

$$\int K(u)\Big[\sum_{j=1}^d \frac{\partial f(x)}{\partial x_j}h_j u_j + \frac{1}{2}\sum_{j,k=1}^d \frac{\partial^2 f(x)}{\partial x_j\partial x_k}h_j h_k u_j u_k + O\Big(\sum_{j=1}^d h_j^3\Big)\Big]\mathrm{d}u,$$

其中用到了估计 $h_jh_kh_l \leqslant \max\{h_j,h_k,h_l\}^3 = \max\{h_j^3,h_k^3,h_l^3\} \leqslant \sum_{j=1}^d h_j^3$ 去处理含有三阶导数的项. 对所有的 j

$$\int u_jK(u)\mathrm{d}u = \prod_{k\neq j}\int L(u_k)\mathrm{d}u_k \cdot \int u_jL(u_j)\mathrm{d}u_j = 0,$$

且对 $j\neq k$

289

$$\int u_ju_kK(u)\mathrm{d}u = \prod_{l\neq j,k}\int L(u_l)\mathrm{d}u_l \cdot \int u_jL(u_j)\mathrm{d}u_j\int u_kL(u_k)\mathrm{d}u_k = 0,$$

同时也有

$$\int u_j^2K(u)\mathrm{d}u = \prod_{l\neq j}\int L(u_l)\mathrm{d}u_l\int u_j^2L(u_j)\mathrm{d}u_j = \widetilde{L}_2.$$

结合起来，就得到

$$E(\hat{f}_T(x))-f(x) = \Big\{\frac{\widetilde{L}_2}{2}\sum_{j=1}^d h_j^2\frac{\partial^2 f}{\partial x_j^2} + O\Big(\sum_{j=1}^d h_j^3\Big)\Big\}.$$

由 $\hat{f}_T(x)$ 的被加项的独立性知

$$\begin{aligned}(Th_1\cdots h_d)\mathrm{Var}(\hat{f}_T(x)) &= \frac{1}{h_1\cdots h_d}\mathrm{Var}K\Big(\frac{X_1-x}{h}\Big)\\ &\leqslant EK^2\Big(\frac{X_1-x}{h}\Big),\end{aligned}$$

其中

$$\begin{aligned}EK^2\Big(\frac{X_1-x}{h}\Big) &= \int K^2\Big(\frac{z-x}{h}\Big)f(z)\mathrm{d}z\\ &= h_1\cdots h_d\int K^2(u)f(x+hu)\mathrm{d}u\\ &= h_1\cdots h_d\int K^2(u)[f(x)+\nabla f(x)'(hu)+\frac{1}{2}(hu)'Df(x^*)(hu)]\mathrm{d}u\end{aligned}$$

$$=h_1\cdots h_d\Big(f(x)\int K^2(u)\mathrm{d}u+\int K^2(u)\nabla f(x)'hu\,\mathrm{d}u$$
$$+\frac{1}{2}\int K^2(u)(hu)'Df(x^*)(hu)\mathrm{d}u\Big),$$

其中 $\int K^2(u)\mathrm{d}u=L_2^d$，$x^*$ 是介于 x 和 $x+hu$ 之间的某个数，$\nabla f(x)$ 为 f 关于 x 的梯度，$Df(x^*)$ 是 f 在 x^* 的二阶偏导数矩阵．记 $\|x\|=\sum_{j=1}^{d}|x_i|$ 为向量 $x=(x_1,\cdots,x_d)'\in\mathbf{R}^d$ 的 l_1 范数．矩阵范数也记为 $\|\cdot\|$．由假定(8.9)和(8.10)得到 $\|Df(x^*)\|\leqslant c<\infty$，结合不等式 $|x'Ax|\leqslant\|A\|\cdot\|x\|^2$ 可导出

$$\int|K^2(u)(hu)'Df(x^*)(hu)|\mathrm{d}u\leqslant c\int K^2(u)\Big(\sum_{j=1}^{d}h_j|u_j|\Big)^2\mathrm{d}u$$
$$\leqslant 2\mathrm{d}c\int K^2(u)\sum_{j=1}^{d}h_j^2u_j^2\mathrm{d}u.$$

其中用到了实数 $\alpha_1,\cdots,\alpha_d$ 的不等式 $\Big(\sum_{j=1}^{d}\alpha_j\Big)^2=O\Big(\sum_{j=1}^{d}\alpha_j^2\Big)$．事实上，由 $\alpha_j\alpha_k\leqslant\alpha_j^2+\alpha_k^2$ 知 $\Big(\sum_{j=1}^{d}\alpha_j\Big)^2=\sum_{j,k=1}^{d}\alpha_j\alpha_k\leqslant 2d\sum_{j=1}^{d}\alpha_j^2$．因为对实数 $\alpha_1,\cdots,\alpha_d,\beta_1,\cdots,\beta_d$，有

$$\sum_{j=1}^{d}\alpha_j^2\beta_j^2\leqslant\sum_{j=1}^{d}\alpha_j^2(\beta_1^2+\cdots+\beta_d^2)\leqslant\Big(\sum_{j=1}^{d}\alpha_j^2\Big)\Big(\sum_{j=1}^{d}\beta_j^2\Big).$$

290

又因为

$$\int u_j^2K^2(u)\mathrm{d}u=\int u_j^2\prod_{k=1}^{d}L(u_k)\mathrm{d}u=\prod_{k\neq j}\int L(u_k)\mathrm{d}u_j\int u_j^2L(u_j)\mathrm{d}u_j<\infty.$$

所以得到

$$\int|K^2(u)(hu)'Df(x^*)(hu)|\mathrm{d}u\leqslant 2dc\int K^2(u)\Big(\sum_{j=1}^{d}h_j^2\Big)\Big(\sum_{j=1}^{d}u_j^2\Big)\mathrm{d}u$$
$$=2dc\int\Big(K^2(u)\sum_{j=1}^{d}u_j^2\Big)\mathrm{d}u\sum_{j=1}^{d}h_j^2$$
$$=2dc\sum_{j=1}^{d}\int K^2(u)u_j^2\mathrm{d}u\sum_{j=1}^{d}h_j^2$$
$$=0\Big(\sum_{j=1}^{d}h_j^2\Big).$$

进一步，

$$\Big|\int K^2(u)\nabla f(x)'hu\,\mathrm{d}u\Big|\leqslant\int K^2(u)\|\nabla f(x)]\|\,\|hu\|\mathrm{d}u$$
$$\leqslant\sqrt{\int K^2(u)\|\nabla f(x)\|^2\mathrm{d}u}\sqrt{\int K^2(u)\|hu\|^2\mathrm{d}u}$$
$$=O\Big(\int K^2(u)\sum_{j=1}^{d}h_j^2u_j^2\mathrm{d}u\Big)$$

$$=O\Big(\sum_{j=1}^{d}h_j^2\Big).$$

因此由满足 $E(Z^2)<\infty$ 的随机变量 Z，有公式 $\mathrm{Var}(Z)=E(Z^2)-(E(Z))^2$ 知

$$\mathrm{Var}K\Big(\frac{X_1-x}{h}\Big)\leqslant EK^2\Big(\frac{X_1-x}{h}\Big)=O\Big(\sum_{j=1}^{d}h_j^2\Big).$$

注意到 $\int K^2(u)\mathrm{d}uf(x)=L_2^d f(x)$，所以得到

$$(Th_1\cdots h_d)\mathrm{Var}(\hat{f}_T(x))=L_2^d f(x)+O\Big(\sum_{j=1}^{d}h_j^2\Big),$$

这就证明了(8.13)式. 均方误差的公式是显然的，因为

291

$$\mathrm{MSE}(\hat{f}_T(x));f(x))=\mathrm{Bias}(\hat{f}_T(x));(\hat{f}_T(x))^2+\mathrm{Var}(\hat{f}_T(x)).$$ ■

得到偏差和方差的显式表达式后，我们可以建立 d 维核密度估计一致性的条件，并且确定可以平衡偏差和方差的最优带宽.

推论 8.5.2 如果

$$\max(h_1,\cdots,h_d)\to 0,\quad Th_1\cdots h_d\to\infty.$$

那么，核密度估计 $\hat{f}_T(x)$ 是一致的. 最优带宽 h_1^*，…，h_d^* 是由 $h_j^4=O((Th_j)^{-1})$ 推导得的

$$h_j=c_jT^{-1/(d+4)},\quad j=1,\cdots,d.$$

值得评论的是：在如此一般条件下核密度是一致的；如果 X_t 坐标间有相依性也是可行的，这种情况前面已经出现过，但是为了简化讨论，在乘积核中假设坐标间是独立的.

为了方便其后的讨论，提及一下下列简单事实：如果选用最优带宽 $h_j=c_jT^{-1/(d+4)}$，则

$$\sqrt{Th_1\cdots h_d}h_j^2=\sqrt{Th_1\cdots h_d h_j^4}=\sqrt{C_jT^{1-d/(d+4)-4/(d+4)}}=\sqrt{C_j},\tag{8.14}$$

其中 $C_j=\prod_{k=1,k\neq j}^{d}c_kc_j^5$ 是常数，所以对于任意给定的 φ_1，…，φ_d，$\sqrt{Th_1\cdots h_d}\sum_{j=1}^{d}h_j^2\varphi_j$ 也是常数，对任意 $\gamma>0$，当 $T\to\infty$ 时，有 $\sqrt{Th_1\cdots h_d}h_j^2h^\gamma\to 0$，且 $\sqrt{Th_1\cdots h_d}\sum_{j=1}^{d}h_j^2h^\gamma\to 0$. 所以有下列重新缩放偏差的结论.

评注 8.5.3 当用最优带宽 $h_j=c_jT^{-1/(d+4)}(j=1,\cdots,d)$ 时，则有

$$\sqrt{Th_1\cdots h_d}\mathrm{Bias}(\hat{f}_T(x))=\frac{\widetilde{L}_2}{2}\sum_{j=1}^{d}\sqrt{C_j}\frac{\partial^2 f}{\partial x_j^2}+o(1).$$

事实上，

$$\begin{aligned}\sqrt{Th_1\cdots h_d}\mathrm{Bias}(\hat{f}_T(x))&=\frac{\widetilde{L}_2}{2}\sum_{j=1}^{d}\sqrt{Th_1\cdots h_d}h_j^2\frac{\partial^2 f}{\partial x_j^2}+O\Big(\sum_{j=1}^{d}\sqrt{Th_1\cdots h_d}h_j^3\Big)\\&=\frac{\widetilde{L}_2}{2}\sum_{j=1}^{d}\sqrt{C_j}\frac{\partial^2 f}{\partial x_j^2}+o(1).\end{aligned}$$

现在对核密度估计量是渐近正态的所需条件进行研究. 将会用到下列形式的中心极限定理，其阵列的每行随机变量是独立同分布的.

定理 8.5.4 设$\{\xi_{Tt}:1\leqslant t\leqslant T,\ T\in\mathbf{N}\}$是随机变量阵列，满足：对每个$T\in\mathbf{N}$，$\xi_{T1}$，…，$\xi_{TT}$是独立同分布的．如果当$T\to\infty$时有

292

$$\mathrm{Var}(\xi_{T1})\to\sigma^2\in(0,\infty),$$

则当$T\to\infty$时有

$$\frac{1}{\sqrt{T}}\sum_{t=1}^{T}[\xi_{Tt}-E(\xi_{Tt})]\xrightarrow{d}N(0,\sigma^2).$$

定理 8.5.5 设X_1，…，X_T是取值于$\mathbf{R}^d$的独立同分布随机向量列，密度函数为$f(x)$，又设f有紧支撑集$\mathrm{supp}(f)$或K有紧支撑集且是有界的，如果对$x\in\mathrm{supp}(f)$，$f(x)$的三阶偏导存在且有界，并且当$T\to\infty$时

$$h_j\to0,j=1,\cdots,d,Th_1\cdots h_d\to\infty$$

和当$T\to\infty$时

$$(Th_1\cdots h_d)\sum_{j=1}^{d}h_j^6\to0,\tag{8.15}$$

则当$T\to\infty$时有

$$\sqrt{Th_1\cdots h_d}\left[\hat{f}_T(x)-f(x)-\frac{\widetilde{L}_2}{2}\sum_{j=1}^{d}h_j^2\frac{\partial^2 f(x)}{\partial x_j^2}\right]\xrightarrow{d}N(0,L_2^d f(x)).$$

证明 已经知道$\hat{f}_T(x)$是$f(x)$的一个有偏估计，考虑分解

$$\hat{f}_T(x)-f(x)=(\hat{f}_T(x)-E\,\hat{f}_T(x))+(E\,\hat{f}_T(x)-f(x)).$$

第一步先证明：当$T\to\infty$时

$$\sqrt{Th_1\cdots h_d}[\hat{f}_T(x)-E\,\hat{f}_T(x)]\xrightarrow{d}N(0,L_2^d f(x)),\tag{8.16}$$

注意到

$$\sqrt{Th_1\cdots h_d}[\hat{f}_T(x)-E\,\hat{f}_T(x)]=\frac{1}{\sqrt{T}}\sum_{t=1}^{T}[\xi_{Tt}-E\xi_{Tt}],$$

且

$$\xi_{Tt}=\frac{1}{\sqrt{h_1\cdots h_d}}K\left(\frac{x-X_t}{h}\right),\quad t=1,\cdots,T.$$

293

首先验证$\mathrm{Var}(\xi_{T1})$是否收敛．因为$x=(x_1,\ \cdots,\ x_d)$，当$T\to\infty$时

$$\begin{aligned}E(\xi_{T1})^2&=\frac{1}{h_1\cdots h_d}\int\cdots\int\prod_{j=1}^{d}L\left(\frac{x_j-z_j}{h_j}\right)^2 f(z_1,\cdots,z_d)\prod_{j=1}^{d}\mathrm{d}z_j\\&=\int\cdots\int\prod_{j=1}^{d}L(u_j)^2 f(x_1+h_1u_1,\cdots,x_d+h_du_d)\prod_{j=1}^{d}\mathrm{d}u_j\\&=\int\cdots\int\prod_{j=1}^{d}L(u_j)^2\left[f(x_1,\cdots,x_d)+O\left(\sum_{j=1}^{d}h_j\right)\right]\prod_{j=1}^{d}\mathrm{d}u_j\\&\to L_2^d f(x),\end{aligned}$$

其中用到了泰勒展开式

$$f(x+hu)=f(x)+\nabla f(x^*)'(hu)=f(x)+O\Big(\sum_{j=1}^d h_j\Big),$$

这里 x^* 是介于 x 和 $x+hu$ 之间的某个数. 进一步，由类似的讨论可知

$$E(\xi_{T1})=O(\sqrt{h_1\cdots h_d}),$$

所以，当 $T\to\infty$ 时

$$\mathrm{Var}(\xi_{T1})\to L_2^d f(x),$$

因此运用定理 8.5.4 即可得(8.16)式.

接下来研究在给定条件下有偏项的性质. 由于用因子 $\sqrt{Th_1\cdots h_d}$ 缩放 $\hat{f}_T(x)-f(x)$，所以要考虑评注 8.5.3 中的归纳和总结. 首先，考虑缩放偏差的第二项. 有

$$\begin{aligned}\sqrt{Th_1\cdots h_d}\sum_{j=1}^d h_j^3 &= \sum_{j=1}^d \sqrt{Th_1\cdots h_d h_j^6}\\ &\leqslant \sum_{j=1}^d \sqrt{Th_1\cdots h_d \sum_{j=1}^d h_j^6}\\ &= d\sqrt{Th_1\cdots h_d \sum_{j=1}^d h_j^6}.\end{aligned}$$

类似地，记 $\varphi_j=\dfrac{\partial^2 f}{\partial x_j^2}$，由第一项得到

294

$$\sqrt{Th_1\cdots h_d}\sum_{j=1}^d h_j^2\varphi_j\leqslant\sqrt{CTh_1\cdots h_d\sum_{j=1}^d h_j^4},$$

其中 $C=d^2\max\limits_{1\leqslant j\leqslant d}\varphi_j^2$. 假定 $Th_1\cdots h_d\sum_{j=1}^d h_j^6=o(1)$ 确保了 $\sqrt{Th_1\cdots h_d}(E\hat{f}_T(x)-f(x))$ 的第二项在 $T\to\infty$ 时收敛于 0. 选择最优带宽使得第一项收敛于一个常数，见评注 8.5.3. 然而，在当前定理的更一般条件下，该项从 $\sqrt{Th_1\cdots h_d}[\hat{f}_T(x)-f(x)]$ 中减去，即用 $\sqrt{Th_1\cdots h_d}[\hat{f}_T(x)-f(x)-b_h(x)]$ 代替，其中，

$$b_h(x)=\frac{\widetilde{L}_2}{2}\sum_{j=1}^d h_j^2\frac{\partial^2 f(x)}{\partial x_j^2}.$$

对应的分解为

$$\begin{aligned}\sqrt{Th_1\cdots h_d}[\hat{f}_T(x)-f(x)-b_h(x)] &= \sqrt{Th_1\cdots h_d}[\hat{f}_T(x)-E\hat{f}_T(x)]\\ &\quad+\sqrt{Th_1\cdots h_d}[E\hat{f}_T(x)-f(x)-b_h(x)],\end{aligned}$$

如上所示有

$$\sqrt{Th_1\cdots h_d}[E\hat{f}_T(x)-f(x)-b_h(x)]=O(\sqrt{Th_1\cdots h_d}\sum_{j=1}^d h_j^3)=o(1).$$

由 Slutzky 引理可得： 当 $T\to\infty$ 时

$$\begin{aligned}\sqrt{Th_1\cdots h_d}[\hat{f}_T(x)-f(x)-b_h(x)] &= \sqrt{Th_1\cdots h_d}[\hat{f}_T(x)-E\hat{f}_T(x)]+o(1)\\ &\xrightarrow{d} N(0,L_2^d f(x)).\end{aligned}$$

■

8.5.2 非参数回归

设(Y, X)，(Y_t, X_t)，$t=1$，…，T，是独立同分布的随机向量，其中Y是单变量的，$X=(X_1, \cdots, X_d)'$是d维随机向量，对应的联合密度函数为$f_{(Y,X)}(y, x)$，其边际密度分别为$f_Y(y)$和$f_X(x)$. 设对某个光滑函数$m(x)$有

$$Y_t = m(X_t) + \varepsilon_t, \quad t = 1,\cdots,T, \tag{8.17}$$

其中，误差项$\{\varepsilon_t\}$是独立同分布的，且满足$E(\varepsilon_t | X_t)=0$，$t=1$，…，T. 在9.2节中将推广误差项为依赖过程. 由此可知$m(X_t)$是给定X时Y的条件均值

$$m(X_t) = E(Y_t | X_t), \quad \text{a.s.}$$

进一步，对于给定的X_t，允许有**条件异方差**并假定

295

$$E(\varepsilon_t^2 | X = X_t) = \sigma^2(X_t),$$

其中，σ^2是回归量的连续函数，满足

$$\int K^2(x)\sigma^2(x) f_X(x)\mathrm{d}x < \infty.$$

众所周知，

$$E(Y|X = x) = \int y f_{Y|X=x}(y)\mathrm{d}y,$$

其中

$$f_{Y|X=x}(y) = \begin{cases} \dfrac{f_{(Y,X)}(y,x)}{f_X(x)}, & y \in \mathbf{R}, x \in \{f_X \neq 0\}, \\ f_Y(y), & y \in \mathbf{R}, x \in \{f_X = 0\}. \end{cases}$$

因为$P(X \in \{f_X \neq 0\})=0$，所以通常忽略后一种情形，并不失一般性假设$f_X>0$. 为了估计$E(Y|X=x)$，只需估计条件密度$f_{Y|X=x}(y)$. 这一问题可以归结为估计联合$(d+1)$维密度$f_{(Y,X)}(y, x)$和$d$维密度$f_X(x)$的问题，这在前面已经讨论过了. 给定估计量$\hat{f}_T(y, x)$和$\hat{f}_T(x)$，可以用

$$\hat{f}_T(y|x) = \hat{f}_{Y|X=x}(y) = \frac{\hat{f}_T(y,x)}{\hat{f}_T(x)}$$

去估计条件密度. 下面用核密度估计量去估计.

选择一个y坐标的光滑核G满足假定(8.9)和(8.10)，带宽$h_0>0$. 定义缩放乘积核

$$\frac{1}{h_0\cdots h_d} K\left(\frac{x}{h}\right) G\left(\frac{y}{h_0}\right) = \frac{1}{h_0\cdots h_d} \prod_{j=1}^{d} L\left(\frac{x_j}{h_j}\right) G\left(\frac{y}{h_0}\right), \quad x \in \mathbf{R}^d, y \in \mathbf{R},$$

其中$x/h=(x_1/h_1, \cdots, x_d/h_d)$. 对应的$(Y, X)$联合密度的估计量为

$$\hat{f}_T(y,x) = \frac{1}{Th_0\cdots h_d} \sum_{t=1}^{T} K\left(\frac{x - X_t}{h}\right) G\left(\frac{y - Y_t}{h_0}\right), \quad x \in \mathbf{R}^d, y \in \mathbf{R}.$$

现计算对应的条件均值

$$\int y\,\hat{f}_T(y|x)\mathrm{d}y = \frac{1}{\hat{f}_T(x)}\int y\,\hat{f}_T(y,x)\mathrm{d}y.$$

则有

296

$$\int y\,\hat{f}_T(y,x)\mathrm{d}y = \frac{1}{Th_0\cdots h_d}\sum_{t=1}^{T}K\Big(\frac{x-X_t}{h}\Big)\int yG\Big(\frac{y-Y_t}{h_0}\Big)\mathrm{d}y,$$

其中

$$\begin{aligned}\int yG(\frac{y-Y_t}{h_0})\mathrm{d}y &= h_0\int(Y_t+h_0u)G(u)\mathrm{d}u \\ &= h_0Y_t + h_0^2\int uG(u)\mathrm{d}u \\ &= h_0Y_t.\end{aligned}$$

因此，得到

$$\int y\,\hat{f}_T(y,x)\mathrm{d}y = \frac{1}{Th_1\cdots h_d}\sum_{t=1}^{T}K\Big(\frac{x-X_t}{h}\Big)Y_t.$$

上述讨论引出了下列定义.

定义 8.5.6 条件期望 $m(x)=E(Y|X=x)$ 的 **Nadaraya-Watson 估计量**定义为

$$\hat{m}_T(x) = \frac{\sum_{t=1}^{T}K\Big(\frac{x-X_t}{h}\Big)Y_t}{\sum_{t=1}^{T}K\Big(\frac{x-X_t}{h}\Big)},\quad x\in\mathbf{R}^d.$$

注意到$\hat{m}_T(x)=\hat{m}_T(x;\ Y_1,\ \cdots,\ Y_T)$是 $Y_1,\ \cdots,\ Y_T$ 的加权平均，其权重之和为 1. 因此，对固定的 x 和任一常数 a，有下列计算法则

$$\hat{m}_T(x;Y_1,\cdots,Y_T)-a = \hat{m}_T(x;Y_1-a,\cdots,Y_T-a). \tag{8.18}$$

下一个目标是研究差$\hat{m}_T(x)-m(x)$的性质. 一个基本的恒等式是

$$\hat{m}_T(x)-m(x) = \frac{[\hat{m}_T(x)-m(x)]\,\hat{f}_T(x)}{\hat{f}_T(x)} = \frac{\hat{g}_T(x)}{\hat{f}_T(x)}. \tag{8.19}$$

例如，如果已经建立了分子的某种缩放形式依分布收敛，或许经过核密度估计中一些必要的偏差修正，$\hat{m}_T$ 的渐近正态性会变得相当容易. 考虑

$$\hat{g}_T(x) = [\hat{m}_T(x)-m(x)]\,\hat{f}_T(x) = \frac{1}{Th_1\cdots h_d}\sum_{t=1}^{T}K\Big(\frac{x-X_t}{h}\Big)(Y_t-m(x)).$$

模型(8.17)，即 $Y_t-m(X_t)=\varepsilon_t$，得到分解

297

$$\hat{g}_T(x) = \tilde{g}_T(x)+\bar{g}_T(x),$$

其中

$$\tilde{g}_T(x) = \frac{1}{Th_1\cdots h_d}\sum_{t=1}^{T}K\Big(\frac{x-X_t}{h}\Big)\varepsilon_t,$$

$$\bar{g}_T(x) = \frac{1}{Th_1\cdots h_d}\sum_{t=1}^{T}K\Big(\frac{x-X_t}{h}\Big)[m(X_t)-m(x)].$$

下面分别讨论这两项. 下列定理表明 $\bar{g}_T(x)$依 L_2 收敛到一个包含带宽的确定表达式.

定理 8.5.7 假设 f 和 m 的三阶偏导存在且有界，K 是有界的、对称的且有紧支集，则

（ⅰ）有

$$E\overline{g}_T(x) = \widetilde{L}_2 \sum_{j=1}^{d} h_j^2 b_j(m;f) + O\Big(\sum_{j=1}^{d} h_j^3\Big),$$

其中

$$b_j(m;f) = \frac{\partial m(x)}{\partial x_j}\frac{\partial f(x)}{\partial x_j} + \frac{1}{2}\frac{\partial^2 m(x)}{\partial x_j^2} f(x). \tag{8.20}$$

（ⅱ）对独立同分布的随机向量(Y_t, X_t)，$t=1, \cdots, T$，

$$\mathrm{Var}(\overline{g}_T(x)) = O\Big(\frac{1}{Th_1\cdots h_d}\sum_{j=1}^{d} h_j^2\Big).$$

证明 为了简化记号，记 $f=f_X$.

（ⅰ）因为

$$\begin{aligned} E\overline{g}_T(x) &= \frac{1}{h_1\cdots h_d}\int K\big(\frac{x-z}{h}\big)[m(z)-m(x)]f(z)\mathrm{d}z \\ &= \int K(u)[m(x+hu)-m(x)]f(x+hu)\mathrm{d}u, \end{aligned}$$

泰勒展开表明上述表达式等于

$$\int K(u)\Big[\nabla m(x)(hu) + \frac{1}{2}(hu)'Dm(x)(hu) + r_m\Big][f(x)+\nabla f(x)(hu)+r_f]\mathrm{d}u,$$

298

由假设(8.9)知

$$f(x)\nabla m(x)\int uK(u)\mathrm{d}u = 0,$$

所以有

$$\int K(u)\Big[\frac{1}{2}(hu)'Dm(x)(hu) + r_m\Big][\nabla f(x)(hu)+r_f]\mathrm{d}u.$$

因此，由 m 和 f 关于高阶偏导的假定，以及被积函数仅在一个紧集上有值，所以

$$r_m = \frac{1}{3!}\sum_{i,j,k=1}^{d}\frac{\partial^3 m(x^*)}{\partial x_i \partial x_j \partial x_k} h_i h_j h_k u_i u_j u_k = O\Big(\sum_{j=1}^{d} h_j^3\Big), \tag{8.21}$$

$$r_f = \frac{1}{2}(hu)'Df(x^+)(hu) = O\Big(\sum_{j=1}^{d} h_j^2\Big), \tag{8.22}$$

其中 x^* 和 x^+ 介于 x 和 $x+hu$ 之间. 现在扩展乘积[…][…]，有

$$\begin{aligned} E\overline{g}_T(x) = &\int K(u)[\nabla m(x)(hu)\nabla f(x)(hu) + \frac{1}{2}(hu)'Dm(x)(hu)f(x)]\mathrm{d}u \\ &+ O\Big(\sum_{j=1}^{d} h_j^3\Big), \end{aligned}$$

其中第一个积分，记为 I_T，其阶为 $O\Big(\sum_{j=1}^{d} h_j^2\Big)$，精确地，有

$$
\begin{aligned}
I_T &= \int K(u) \sum_{j,k=1}^{d} \left[\frac{\partial m(x)}{\partial x_j} \frac{\partial f(x)}{\partial x_k} + \frac{1}{2} \frac{\partial^2 m(x)}{\partial x_j \partial x_k} f(x) \right] h_j h_k u_j u_k \mathrm{d}u \\
&= \int K(u) \sum_{j=1}^{d} \left[\frac{\partial m(x)}{\partial x_j} \frac{\partial f(x)}{\partial x_j} + \frac{1}{2} \frac{\partial^2 m(x)}{\partial x_j^2} f(x) \right] h_j^2 u_j^2 \mathrm{d}u \\
&= \widetilde{L}_2 \sum_{j=1}^{d} h_j^2 b_j(m; f),
\end{aligned}
$$

其中 $b_j(m; f)$由(8.20)式给出，其余项的阶为 $O\Big(\sum_{j=1}^{d} h_j^3\Big)$，实际上有的项为 0，比如

299

$$
\int K(u) \nabla m(x)'(hu) f(x) \mathrm{d}u = f(x) \nabla m(x)' h \Big(\int K(u) u_j \mathrm{d}u \Big)_{j=1}^{d} = 0,
$$

又比如，

$$
\begin{aligned}
J_h &= \int K(u) \left[\frac{1}{2} (hu)' D m(x) (hu) \nabla f(x) (hu) \right] \mathrm{d}u \\
&= \frac{1}{2} \int K(u) \sum_{i,j,k=1}^{d} \frac{\partial^2 m(x)}{\partial x_j \partial x_k} \frac{\partial f(x)}{\partial x_i} h_i h_j h_k u_i u_j u_k \mathrm{d}u \\
&= \frac{1}{2} \int K(u) \sum_{j=1}^{d} \frac{\partial^2 m(x)}{\partial x_j^2} \frac{\partial f(x)}{\partial x_j} h_j^3 u_j^3 \mathrm{d}u \\
&= O\Big(\sum_{j=1}^{d} h_j^3 \Big).
\end{aligned}
$$

（ⅱ）由独立同分布的假定以及常用的代换 $u_j=(z_j-x_j)/h_j$，$j=1$，…，d，得到

$$
\begin{aligned}
\mathrm{Var}(\bar{g}_T) &= \frac{1}{T h_1^2 \cdots h_d^2} \mathrm{Var}\Big(K\Big(\frac{x-X_t}{h}\Big)(m(X_T)-m(x))\Big) \\
&\leqslant \frac{1}{T h_1^2 \cdots h_d^2} E\Big(K\Big(\frac{x-X_t}{h}\Big)^2 (m(X_T)-m(x))^2\Big) \\
&= \frac{1}{T h_1 \cdots h_d} \int K^2(u) [m(x+hu)-m(x)]^2 f(x+hu) \mathrm{d}u.
\end{aligned}
$$

后一个表达式类似于（ⅰ）的处理，由此得到估计

$$
\mathrm{Var}(\bar{g}_T) = \frac{1}{T h_1 \cdots h_d} O\Big(\sum_{j=1}^{d} h_j^2 \Big).
$$

注意，定理 8.5.7 的证明并没有假定 $m(x)$是条件均值 $E(Y|X=x)$. 事实上，该结论可以推广到更为一般的函数，从而得到下列平滑平均的弱大数定律. ■

定理 8.5.8（核平滑的弱大数定律） 假设 $X_t \overset{\text{i.i.d}}{\sim} f(x)$，$\psi(z, x)$是定义在 $X(\Omega)\times X(\Omega)$上的函数，且满足 $E|\psi(X_1, x)|^2<\infty$. 若 f 和 ψ 的三阶偏导存在且有界，则

300

$$
\begin{aligned}
&\frac{1}{T h_1 \cdots h_d} \sum_{t=1}^{T} K\Big(\frac{X_t-x}{h}\Big)[\psi(X_t,x)-\psi(x,x)] \\
&= \int K\Big(\frac{z-x}{h}\Big)[\psi(z,x)-\psi(x,x)](z)\mathrm{d}z + o_P(1),
\end{aligned}
\tag{8.23}
$$

且

$$\int K\left(\frac{z-x}{h}\right)[\psi(z,x)-\psi(x,x)](z)\mathrm{d}z=o(1).$$

定理 8.5.9 设 f_X 和 m 的三阶偏导数存在且有界，K 有界且有紧支集，如果

$$\max_{j=1,\cdots,d} h_j \to 0, \quad Th_1\cdots h_d \to \infty$$

且当 $T\to\infty$ 时，

$$(Th_1\cdots h_d)\sum_{j=1}^d h_j^6 \to 0, \tag{8.24}$$

则当 $T\to\infty$ 时，有

$$\sqrt{Th_1\cdots h_d}\left(\hat{m}_T(x)-m(x)-\widetilde{L}_2\sum_{j=1}^d h_j^2 b_j(m;f)\right)\xrightarrow{d} N(0,L_2^d\sigma^2(x)/f_X(x))$$

对任何满足 $f_X(x)>0$ 的 x 成立.

证明 由分解(8.19)式知

$$\hat{f}_T(x)\sqrt{Th_1\cdots h_d}\left(\hat{m}_T(x)-m(x)-L_2\sum_{j=1}^d h_j^2 b_j(m;f)\right)=\hat{f}_T(x)(A_T(x)+B_T(x)),$$

其中

$$A_T(x)=\sqrt{Th_1\cdots h_d}\,\widetilde{g}_T(x),$$

$$B_T(x)=\sqrt{Th_1\cdots h_d}\left(\bar{g}_T(x)-\widetilde{L}_2\sum_{j=1}^d h_j^2 b_j(m;f)\right).$$

根据之前的命题结果

$$EB_T(x)=O\left(\sqrt{Th_1\cdots h_d}\sum_{j=1}^d h_j^3\right),$$

且

$$\mathrm{Var}B_T(x)=O\left(\sum_{j=1}^d h_j^2\right).$$

由假定(8.24)知，当 $T\to\infty$ 时，上述两式的右边均收敛于 0，具体证明参考定理 8.5.5 的证明．这表明：当 $T\to\infty$ 时

$$B_T \xrightarrow{L_2,P} 0.$$

301

考虑

$$A_T(x)=\frac{1}{\sqrt{Th_1\cdots h_d}}\sum_{t=1}^T K\left(\frac{x-X_t}{h}\right)\varepsilon_t$$
$$=\frac{1}{\sqrt{T}}\sum_{t=1}^T \xi_{Tt},$$

其中，对 $t=1,\cdots,T$ 且 $T\geqslant 1$，

$$\xi_{Tt}=\frac{1}{\sqrt{h_1\cdots h_d}}K\left(\frac{x-X_t}{h}\right)\varepsilon_t,$$

显然，被加数是中心化的，由 $E(\varepsilon_t^2\,|\,X_t)=\sigma^2(X_t)$ 可得

$$\mathrm{Var}(\xi_{Tt})=E(E(\xi_{Tt}^2\,|\,X_t))$$

$$=E_{X_t}\left(\frac{1}{h_1\cdots h_d}K^2\left(\frac{x-X_t}{h}\right)\sigma^2(X_t)\right)$$

$$=\int\left(\frac{1}{h_1\cdots h_d}K^2\left(\frac{x-z}{h}\right)\sigma^2(z)f_X(z)\right)\mathrm{d}z$$

$$=\int K^2(u)\sigma^2(x+hu)f_X(x+hu)\mathrm{d}u.$$

显然最后一个表达式收敛于 $L_2^d\sigma^2(x)f_X(x)$. 因此，当 $T\to\infty$时

$$A_T(x)\xrightarrow{d}N(0,L_2^d\sigma^2(x)f_X(x)).$$

由 Slutzky 定理知，当 $T\to\infty$时

$$\hat{f}_T(x)\sqrt{Th_1\cdots h_d}\left(\hat{m}_T(x)-m(x)-\widetilde{L}_2\sum_{j=1}^d h_j^2 b_j(m;f)\right)\xrightarrow{d}N(0,L_2^d\sigma^2(x)f_X(x)),$$

在条件$\hat{f}_T(x)\xrightarrow{P}f(x)$，$T\to\infty$下，由 Slutzky 引理知定理结论成立. ■

8.6 线性过程的中心极限定理

前面已经看到，线性过程是一类重要的时间序列. 因此，本节详细讨论此类过程的中心极限定理. 使用 Beveridge-Nelson 分解，将得到中心极限定理的一个优美且通俗易懂的证明.

定义和定理 8.6.1 设

302

$$X_t=\sum_{i=0}^{\infty}\alpha_i\varepsilon_{t-i},\quad t\in\mathbf{Z}$$

是一个线性过程，其系数$\{\alpha_i\}$满足 $\sum_{i=0}^{\infty}|\alpha_i|<\infty$，则分解

$$X_t=\bar{\alpha}\varepsilon_t+\widetilde{\varepsilon}_{t-1}-\widetilde{\varepsilon}_t,\quad t\in\mathbf{Z}$$

称为 **Beveridge-Nelson 分解**，其中 $\bar{\alpha}=\sum_{i=0}^{\infty}\alpha_i$，且

$$\widetilde{\varepsilon}_t=\sum_{i=0}^{\infty}\widetilde{\alpha}_i\varepsilon_{t-i},\quad \widetilde{\alpha}_i=\sum_{k=i+1}^{\infty}\alpha_k,$$

且线性过程 $X_t=\alpha(L)\varepsilon_t$ 可写为

$$X_t=\alpha(1)\varepsilon_t-(1-L)\widetilde{\varepsilon}_t,\quad \widetilde{\varepsilon}_t=\widetilde{\alpha}(L)\varepsilon_t,$$

其中滞后算子$\widetilde{\alpha}(L)$的系数为$\widetilde{\alpha_i}$.

证明 注意到，对一个系数为$\{\widetilde{\alpha_i}\}$的线性过程

$$\widetilde{\varepsilon}_t=\sum_{i=0}^{\infty}\widetilde{\alpha}_i\varepsilon_{t-i}$$

有

$$-\widetilde{\varepsilon}_t+\widetilde{\varepsilon}_{t-1}=-(\widetilde{\alpha}_0\varepsilon_t+\widetilde{\alpha}_1\varepsilon_{t-1}+\widetilde{\alpha}_2\varepsilon_{t-1}+\cdots)$$
$$+\widetilde{\alpha}_0\varepsilon_{t-1}+\widetilde{\alpha}_1\varepsilon_{t-2}+\cdots,$$

从而得到

$$-\widetilde{\varepsilon}_t+\widetilde{\varepsilon}_{t-1}=-\widetilde{\alpha}_0\varepsilon_t+(\widetilde{\alpha}_0-\widetilde{\alpha}_1)\varepsilon_{t-1}+(\widetilde{\alpha}_1-\widetilde{\alpha}_2)\varepsilon_{t-2}+\cdots$$

特别地，选取

$$\widetilde{\alpha}_i=\sum_{k>i}\alpha_k,$$

则有

$$\widetilde{\alpha}_0=\sum_{k=1}^{\infty}\alpha_k,\quad \widetilde{\alpha}_i-\widetilde{\alpha}_{i+1}=\Big(\sum_{k=i+1}^{\infty}-\sum_{k=i+2}^{\infty}\Big)\alpha_k=\alpha_{i+1}.$$

因此，

$$-\widetilde{\varepsilon}_t+\widetilde{\varepsilon}_{t-1}=-\Big(\sum_{k=1}^{\infty}\alpha_k\Big)\varepsilon_t+\sum_{i=1}^{\infty}\alpha_i\varepsilon_{t-i}.$$

上式两边同时加 $\alpha(1)\varepsilon_t=\Big(\sum_{k=0}^{\infty}\alpha_k\Big)\varepsilon_t$，则得到

$$\alpha(1)\varepsilon_t-\widetilde{\varepsilon}_t+\widetilde{\varepsilon}_{t-1}=\Big(\sum_{k=0}^{\infty}\alpha_k-\sum_{k=1}^{\infty}\alpha_k\Big)\varepsilon_t+\sum_{i=1}^{\infty}\alpha_i\varepsilon_{t-i}=X_t.$$ ■ 303

为了进一步研究，需要下列简单事实.

引理 8.6.2

（ⅰ）若 $\sum_{i=0}^{\infty}i|\alpha_i|<\infty$，则 $\sum_{i=0}^{\infty}i^2\alpha_i^2<\infty$.

（ⅱ）若 $\sum_{i=0}^{\infty}i|\alpha_i|<\infty$，则 Beveridge-Nelson 分解中的过程$\{\widetilde{\varepsilon}_t\}$是弱平稳的(因此是一个 L_2 序列).

证明

（ⅰ）因为存在 $i_0\in\mathbf{N}$ 使得当 $i>i_0$ 时，有 $i|\alpha_i|<1$ 成立. 从而对这些 i，有 $i^2\alpha_i^2<i|\alpha_i|$. 因此

$$\begin{aligned}\sum_{i=0}^{\infty}i^2\alpha_i^2&=\sum_{i=0}^{i_0}i^2\alpha_i^2+\sum_{i=i_0+1}^{\infty}i^2\alpha_i^2\\&<\sum_{i=0}^{i_0}i^2\alpha_i^2+\sum_{i=i_0+1}^{\infty}i|\alpha_i|.\end{aligned}$$

第一项求和只有有限项，因此它是有限的. 第二项求和也有界，其界为 $\sum_{i=0}^{\infty}i|\alpha_i|<\infty$.

（ⅱ）如能证明 $\sum_{i=0}^{\infty}\widetilde{\alpha}_i^2<\infty$,，则有 $\sum_{i=0}^{\infty}|\widetilde{\alpha}_i|<\infty$，从而可以由定理 A. 5. 1 得到（ⅱ）的结论成立. 事实上，因为

$$\begin{aligned}\sum_{i=0}^{\infty}|\widetilde{\alpha}_i|&\leqslant\sum_{i=0}^{\infty}\sum_{j=i+1}^{\infty}|\alpha_j|\\&=\sum_{j=0}^{\infty}\sum_{i=0}^{j-1}|\alpha_j|\\&=\sum_{j=1}^{\infty}j|\alpha_j|<\infty.\end{aligned}$$

所以$\{\widetilde{\alpha_j}\}$是一个收敛于零的序列，从而能找到一个整数 i_0，使得当 $i>i_0$ 时有$|\widetilde{\alpha_i}|\leqslant 1$. 且

$$\sum_{i=0}^{\infty}\widetilde{\alpha}_i^2=\sum_{i=0}^{i_0}\widetilde{\alpha}_i^2+\sum_{i=i_0+1}^{\infty}\widetilde{\alpha}_i^2$$

$$\leqslant\sum_{i=0}^{i_0}\widetilde{\alpha}_i^2+\sum_{i=i_0+1}^{\infty}\widetilde{\alpha}_i^2$$

$$\leqslant\sum_{i=0}^{i_0}\widetilde{\alpha}_i^2+\sum_{i=i_0+1}^{\infty}|\widetilde{\alpha}_i|<\infty.$$ ■

304

下面引入中心极限定理.

定理 8.6.3(线性过程的中心极限定理) 设

$$X_t=\sum_{i=0}^{\infty}\alpha_i\varepsilon_{t-i}$$

是一个线性过程，满足 $\varepsilon_t\overset{\text{i.i.d}}{\sim}(0,\sigma^2)$，其中 $\sigma^2\in(0,\infty)$. 若

$$\sum_{i=0}^{\infty}i|\alpha_i|<\infty,\quad \sum_{i=0}^{\infty}\alpha_i\neq 0,$$

则当 $T\to\infty$ 时，

$$S_T=\frac{1}{\sqrt{T}}\sum_{t=1}^{T}X_t\xrightarrow{d}N(0,\sigma_X^2),$$

其中 $\sigma_X^2=\sigma\left(\sum_{i=0}^{\infty}\alpha_i\right)^2$.

证明 由 Beveridge-Nelson 分解知

$$S_T=U_T+R_T,$$

其中

$$U_T=\frac{\bar{\alpha}}{\sqrt{T}}\sum_{t=1}^{T}\varepsilon_t,$$

$$R_T=\frac{1}{\sqrt{T}}\sum_{t=1}^{T}(\widetilde{\varepsilon}_{t-1}-\widetilde{\varepsilon}_t).$$

显然，

$$R_T=\frac{1}{\sqrt{T}}(\widetilde{\varepsilon}_0-\widetilde{\varepsilon}_T).$$

由满足 $E(\varepsilon_1)=0$ 和 $E(\varepsilon_1^2)<\infty$的独立同分布随机序列$\{\varepsilon_t\}$的经典中心极限定理知：当 $T\to\infty$时

$$U_T\xrightarrow{d}N(0,\bar{\alpha}^2\sigma^2),$$

余下来仅需证明：当 $T\to\infty$时

$$R_T\xrightarrow{P}0.$$

注意到

$$|R_T| \leqslant \frac{1}{\sqrt{T}} |\widetilde{\varepsilon}_0 - \widetilde{\varepsilon}_T| \leqslant \frac{2}{\sqrt{T}} \max_{1 \leqslant t \leqslant T} |\widetilde{\varepsilon}_t|.$$

下面将用到等价性 305

$$\max_{1 \leqslant t \leqslant T} |\widetilde{\varepsilon}_t| > \delta \Leftrightarrow \sum_{t=1}^{T} \widetilde{\varepsilon}_t^2 \mathbf{1}(|\widetilde{\varepsilon}_t| > \delta) > \delta^2. \tag{8.25}$$

对任意 $\delta>0$，其证明如下：首先，如果 $\max_{1 \leqslant t \leqslant T} |\widetilde{\varepsilon_t}| > \delta$，则存在某个 j 使得 $|\widetilde{\varepsilon_j}|>\delta$，则有

$$\sum_{t=1}^{T} \widetilde{\varepsilon}_t^2 \mathbf{1}(|\widetilde{\varepsilon}_t| > \delta) \geqslant |\widetilde{\varepsilon}_j^2| > \delta^2.$$

反之，如果 $\sum_{t=1}^{T} \widetilde{\varepsilon}_t^2 1(|\widetilde{\varepsilon}_t| > \delta) > \delta^2$，则存在某个 j 使得 $|\widetilde{\varepsilon_j}|>\delta$，否则和为 0. 因此 $\max_{1 \leqslant t \leqslant T} |\widetilde{\varepsilon_t}| > \delta$，这就证明了(8.25). 现用 $\widetilde{\varepsilon_t}/\sqrt{T}$ 代替 $\widetilde{\varepsilon_t}$，重复上述步骤，则得到

$$\max_{1 \leqslant t \leqslant T} \frac{|\widetilde{\varepsilon}_t|}{\sqrt{T}} \xrightarrow{P} 0 \Leftrightarrow \frac{1}{T} \sum_{t=1}^{T} \widetilde{\varepsilon}_t^2 \mathbf{1}(|\widetilde{\varepsilon}_t| > \delta \sqrt{T}) \xrightarrow{P} 0.$$

因为 $\{\varepsilon_t\}$ 是一个严平稳的 L_2 过程，所以有：当 $T\to\infty$ 时

$$\begin{aligned} E\left(\frac{1}{T} \sum_{t=1}^{T} \widetilde{\varepsilon}_t^2 \mathbf{1}(|\widetilde{\varepsilon}_t| > \delta \sqrt{T}\right) &= E(\widetilde{\varepsilon}_1^2 \mathbf{1}(|\widetilde{\varepsilon}_1| > \delta \sqrt{T})) \\ &= \int_{-\infty}^{-\delta\sqrt{T}} x^2 \mathrm{d}F(x) + \int_{\delta\sqrt{T}}^{-\infty} x^2 \mathrm{d}F(x) \\ &\to 0, \end{aligned}$$

其中 $F(x)=P(\varepsilon_1 \leqslant x)$，$x \in \mathbf{R}$，因此，当 $T\to\infty$ 时，$R_T \xrightarrow{P} 0$. ■

8.7 混合过程

8.7.1 混合系数

设 $\{X_t\}=\{X_t: t\in Z\}$ 是一个时间序列，记当前时刻为 t. 想知道之前的观测值 X_s，$s<t$ 对现在和未来的值影响有多大，一种方法是观察如下形式的随机变量间的最大相关性

$$Y = f(X_s, X_{s-1}, \cdots), Z = g(X_t, X_{t-1}, \cdots), \tag{8.26}$$

其中 f 和 g 是使得相关性存在的可测函数. 显然，Y 是 $\mathcal{F}_{-\infty}^{s}$ 可测的，Z 是 $\mathcal{F}_t^{\infty}$ 可测的，其中，对 $-\infty \leqslant a \leqslant b \leqslant \infty$

$$\mathcal{F}_a^b = \sigma(X_a, \cdots, X_b). \tag{8.27}$$

反之，任意 $\mathcal{F}_{-\infty}^{s}$ 可测的随机变量 Y 和任意 $\mathcal{F}_t^{\infty}$ 可测的随机变量 Z 都可以写成(8.26)的形式. 为了把这一粗略的描述变成严格的定义，记基本概率空间为$(\Omega, \mathcal{F}, P)$，对某个子 σ 代数 A，记 $L_2(\Omega, A, P)$ 为 Ω 上 $\mathcal{A}$ 可测的随机变量构成的 L_2 空间. 306

定义 8.7.1(ρ 混合) （i）设 $\mathcal{A}$ 和 $\mathcal{B}$ 是 $\mathcal{F}$ 的两个子 σ 代数. 则称

$$\rho(\mathcal{A}, \mathcal{B}) = \sup\{|\mathrm{Cov}(X, Y)| : X \in L_2(\Omega, \mathcal{A}, P), Y \in L_2(\Omega, \mathcal{B}, P)\}$$

为$\mathcal{A}$和$\mathcal{B}$之间的**ρ 混合**系数.

(ⅱ) 设$\{X_t\}$是一个 L_2 序列，则称

$$\rho(k)=\sup_t \rho(\mathcal{F}_{-\infty}^{t-k},\mathcal{F}_t^{\infty}),k\in\mathbf{N}_0$$

为$\{X_t\}$的**滞后 k 期的 ρ 混合系数**，其中$\mathcal{F}_a^b$ 由(8.27)式所给.

(ⅲ) 称一个 L_2 时间序列$\{X_t\}$为 ρ 混合的，如果

$$\lim_{k\to\infty}\rho(k)=0.$$

可以证明

$$\rho(\mathcal{A},\mathcal{B})=\sup\{|\mathrm{Cov}(X,Y)|\},$$

其中上确界为取遍满足下列条件的随机变量 X 和 Y：

$$X\in L_2(\Omega,\mathcal{A},P),Y\in L_2(\Omega,\mathcal{B},P),E(X)=E(Y)=0,\|X\|_{L_2}=\|Y\|_{L_2}=1.$$

进一步，用相关性代替协方差，人们可以去掉二阶矩的限制.

设 $A\in\mathcal{F}_{-\infty}^{t-k}$是一个事件，依赖于直到 $t-k$ 的过去值…，X_{t-k-1}，X_{t-k}. $B\in\mathcal{F}_t^{\infty}$ 是一个事件，依赖于未来的值 X_t，X_{t+1}，…，这意味着在 A 与 B 之间存在 k 个时间单位的时滞. 当滞后时间充分大时，时间序列与滞后值的相关性会消失，这是因为对充分大的滞后值，A 与 B 是下列意义

$$P(A\cap B)\approx P(A)P(B)$$

下渐近独立的. 在最简单的情形下，对足够大的 k，A 和 B 是独立的.

定义 8.7.2　时间序列$\{X_t\}$称为 m 相依的，如果对所有的 $k>m$，$A\in\mathcal{F}_{-\infty}^{t-k}$与 $B\in\mathcal{F}_t^{\infty}$是独立的.

换句话说，称 X_t 是 m 相依的，如果它的任意有限维边际$(X_{s_1}$，…，$X_{s_k})$和$(X_{t_1}$，…，$X_{t_l})$，$s_1<\cdots<s_l\leqslant t_1<\cdots<t_l$ 是独立的，其中 $t_1-s_l>m$. 特别地，一个平稳序列是 0-相依的，当且仅当它是独立同分布的.

307

根据前面的讨论，背离独立性的一个有意义的度量是

$$|P(A\cap B)-P(A)P(B)|,$$

这就产生下面的定义.

定义 8.7.3(α 混合)

(ⅰ) 设$\mathcal{A}$，$\mathcal{B}\subset\mathcal{F}$是子 σ 代数. 称

$$\alpha(\mathcal{A},\mathcal{B})=\sup(|P(A\cap B)-P(A)P(B)|:A\in\mathcal{A},B\in\mathcal{B})$$

为 **α 混合系数**.

(ⅱ) 对时间序列$\{X_t\}$，称

$$\alpha(k)=\sup_t \alpha(\mathcal{F}_{-\infty}^{t-k},\mathcal{F}_t^{\infty})$$

为$\{X_t\}$的**滞后 $k\in\mathbf{N}_0$ 期的强或 α 混合系数**.

(ⅲ) 称一个时间序列$\{X_t\}$为**强或 α 混合的**，如果

$$\lim_{k\to\infty}\alpha(k)=0.$$

注意，这个定义同样适用于非平稳的时间序列，并且，与 ρ 混合相反，它不需要二阶矩存在. 在二阶矩不存在时，ρ 混合仍然有定义，因为它只与随机变量 Y(关于$\mathcal{F}_{-\infty}^{t-k}$可测的)和 Z(关

于 F_t^∞ 可测的)的协方差 $\mathrm{Cov}(Y,Z)$ 有关，而 Y 与 Z 的二阶矩是有限的. 然而此时不能得到关于观测值 X_t 本身的任何有意义的结论，因为不能简单地在定义中取 $Y=X_t$，$Z=X_{t+k}$.

下面将会看到，通常需要对 α-混合系数的序列 $\{\alpha(k): k\in\mathbf{N}_0\}$ 的衰减(率)作一定的假定. 特别地，混合性质本身是不足以得到样本均值的一些重要结论，比如大数定律或中心极限定理.

8.7.2 不等式

本小节归纳总结混合系数的一些基本不等式. 特别地，读者将会看到如何由 α 混合系数来估计协方差.

首先，是一个富有启发性的引理，它指出：ρ 混合总是强于 α 混合.

引理 8.7.4 对 $\mathcal{F}$ 的任意子 σ 代数 $\mathcal{A}$，$\mathcal{B}$，有

$$4\alpha(\mathcal{A},\mathcal{B})\leqslant\rho(\mathcal{A},\mathcal{B}).$$

证明 设 $A\in\mathcal{A}$ 且 $B\in\mathcal{B}$，记 $X=I_A$，$Y=I_B$. 则 $E(X)=P(A)$，$E(Y)=P(B)$，$E(XY)=P(A\cap B)$，且

$$|\mathrm{Cov}(X,Y)|=|P(A\cap B)-P(A)P(B)|.$$

308

因为 $\mathrm{Var}(X)=P(A)(1-P(A))\leqslant 1/4$，$|\mathrm{Cov}(X,Y)|=|\mathrm{Cor}(X,Y)|\sqrt{\mathrm{Var}(X)\mathrm{Var}(Y)}$，所以有下列不等式

$$|P(A\cap B)-P(A)P(B)|\leqslant(1/4)|\mathrm{Cor}(X,Y)|,$$

从而得到

$$4|P(A\cap B)-P(A)P(B)|\leqslant\sup\{|\mathrm{Cor}(X,Y)|:X,Y\in L_2\}.$$

由于右边的式子不依赖于 A 和 B，因此可以去掉 sup，所以得到 $4\alpha(\mathcal{A},\mathcal{B})\leqslant\rho(\mathcal{A},\mathcal{B})$. ■

在处理 α 混合过程时，下面的不等式是至关重要的，其证明相当有技巧，但可以在相关文献中找到. 它表明使用 α 混合系数可以确定一个 $\mathcal{A}$ 可测随机变量 X 和一个 $\mathcal{B}$ 可测随机变量 Y 的协方差的界，即交叉矩 $E(XY)$ 与 $E(X)E(Y)$(假设不独立)的差的绝对值的界，其中假定 X 与 Y 是独立的.

命题 8.7.5 设 p，q，$r\in(1,\infty]$，满足 $\frac{1}{p}+\frac{1}{q}+\frac{1}{r}=1$，$X$ 是一个 $\mathcal{A}$ 可测的随机变量，Y 是一个 $\mathcal{B}$ 可测的随机变量，满足

$$E|X|^p<\infty,\quad E|Y|^q<\infty.$$

则

$$|\mathrm{Cov}(X,Y)|=|E(XY)-E(X)E(Y)|\leqslant 8\|X\|_p\|Y\|_q\alpha(\mathcal{A},\mathcal{B})^{1/r},$$

其中，如果 $x>0$，则 $x^0=1$ 且 $0^0=0$.

需要强调的是：上述不等式必须在 $\frac{1}{p}+\frac{1}{q}$ 小于 1 时才成立. 因此，不能取 $p=q=2$，即不能假定 X 和 Y 的二阶矩有限. 但是，如果存在某个 $\delta>0$，使得 X 和 Y 的 $2+\delta$ 阶绝对矩是存在的，则该不等式是成立的. 此时 $\frac{1}{r}=1-\frac{1}{p}-\frac{1}{q}=\frac{2}{2+\delta}$. 在这种情况下，命中 8.7.5 有下列形式.

命题 8.7.6 设 X 是 $\mathcal{A}$ 可测的，Y 是 $\mathcal{B}$ 可测的．如存在某个 $\delta>0$，使得

$$E|X|^{2+\delta}<\infty, E|Y|^{2+\delta}<\infty,$$

则

$$|\mathrm{Cov}(X,Y)| \leqslant 8\alpha(\mathcal{A},\mathcal{B})^{2/(2+\delta)}\|X\|_{2+\delta}\|Y\|_{2+\delta}.$$ ■

如果 X_t 是独立同分布的，且 $E(X_1)=0$，X_t 的二阶矩有限，由初等的计算知和 $S_n=\sum_{i=1}^{n} X_i$ 满足：矩有界，且

$$E(S_n^2)=O(n).$$

如果 $E(X_1^4)<\infty$成立，则

309

$$E(S_n^4)=O(n^2).$$

对混合过程，要这些阶性质仍然成立，前提是混合系数 $\alpha(k)$，$k\geqslant 0$ 衰减得足够地快，即

$$A_r(\delta)=\sum_{i=0}^{\infty}(i+1)^{r/2-1}\alpha(i)^{\delta/(r+\delta)}<\infty,$$

且序列是 r 阶弱平稳的，即存在 $\delta>0$，其直至 $r+\delta$ 阶矩均存在，且是时间不变的，即

$$E(X_{t_1}\cdots X_{t_r})=E(X_{t_1+h}\cdots X_{t_r+h})$$

对所有的 t_1，…，t_r 和 h 成立．下面详细讨论 $r\in[2, 4]$的重要情形．

定理 8.7.7

（ⅰ）设$\{X_t\}$是弱平稳的，且满足 $E|X_1|^{2+\delta}<\infty$，其中 $\delta>0$，则

$$E(S_n^2)\leqslant 2!8A_2(\delta)\|X_1\|_{2+\delta}^2.$$

（ⅱ）设$\{X_t\}$是 4 阶弱平稳的，且满足 $E|X_1|^{4+\delta}<\infty$．则

$$E(S_n^4)\leqslant 84!A_4(\delta)\|X_1\|_{4+\delta}^4 n^2.$$

证明 对 $r=2$，4 有

$$E(S_n^r)=\sum_{t_1=1}^{n}\cdots\sum_{t_r=1}^{n}E(X_{t_1}\cdots X_{t_r}).$$

注意到，$E(X_{t_1}\cdots X_{t_r})$的所有 $r!$ 项，每一项的下标是有序值 $t_{(1)}\leqslant\cdots\leqslant t_{(r)}$ 的一个排列．由平稳性知，$E(X_{t_1}\cdots X_{t_r})=E(X_0X_{t_2-t_1}\cdots X_{t_r-t_1})$，因此，

$$\begin{aligned}E(S_n^r)&=r!\sum_{1\leqslant t_1\leqslant\cdots\leqslant t_r\leqslant n}E(X_{t_1}\cdots X_{t_r})\\&=r!n\sum_{0\leqslant t_2\leqslant\cdots\leqslant t_r\leqslant n-1}E(X_0X_{t_2}\cdots X_{t_{r-1}})\\&=r!n\sum_{1\leqslant t_1\leqslant\cdots\leqslant t_r-1\leqslant n}E(X_1X_{t_1}\cdots X_{t_{r-1}}).\end{aligned}$$

如 $r=2$，则必须估计 $\sum_{t_1=1}^{n}|E(X_1X_{t_1})|$，可以断言它不大于 $8\|X_1\|_{2+\delta}^2A_2(\delta)$，这就证明了（ⅰ）．为了验证这一断言，注意到 $E(X_1)=0$，所以

310

$$|E(X_0X_j)|=|\mathrm{Cov}(X_0,X_j)|\leqslant 8\|X_1\|_{2+\delta}^2\alpha(j)^{2/(2+\delta)}.$$

由此推出

$$\sum_{t_1=1}^{n}|E(X_1X_{t_1})|\leqslant 8\|X_1\|_{2+\delta}^2\sum_{j=0}^{\infty}\alpha(j)^{2/(2+\delta)}$$
$$=8\|X_1\|_{2+\delta}^2A_2(\delta).$$

$r=4$ 的情形更为复杂且需要适当地重新排列这些项. 记 $\sum\limits_{n,j}^{(h)}$ 为对所有满足下列条件的指标 $1\leqslant t_1\leqslant\cdots\leqslant t_j\leqslant n$ 求和，t_1，…，t_j 中每一指标与它前一指标的差的最大值在 h 取到，即

$$r_h=t_h-t_{h-1}=\max(t_j-t_{j-1},\cdots,t_2-t_1,t_1-t_0),$$

为了简化讨论，令 $t_0=0$，并记 r_h 为最大差值. 现对 $1\leqslant t_1\leqslant t_2\leqslant t_3\leqslant n$ 求和. 首先对位置 $h=1$，2，3 求和，然后对每个 h，求上面定义的有序指标和. 因此，考虑

$$E(S_n^4)=4!n\sum_{h=1}^{3}\sum_{n,3}^{(h)}E(X_1X_{t_1}X_{t_2}X_{t_3}).$$

的右边项. 下面证明

$$\sum_{n,3}^{(h)}|E(X_1X_{t_1}X_{t_2}X_{t_3})|\leqslant 8\|X_1\|_{4+\delta}^4\sum_{r=0}^{n-1}C_r\alpha(r)^{\delta/(4+\delta)}.$$

其中，C_r 是由三元组$(t_1，t_2，t_3)$确定的数，三元组满足 $1=t_0\leqslant t_1\leqslant t_2\leqslant t_3\leqslant n$，$t_h-t_{h-1}=\max\limits_i(t_i-t_{i-1})=l$，$l$ 是一个固定的数，h 为 1，2 或 3. 记对应的情形为

$$1\overset{\ell}{—}t_1t_2t_3,\quad 1t_1\overset{\ell}{—}t_2t_3,\quad 1t_1t_2\overset{\ell}{—}t_3,$$

其中 $\overset{\ell}{—}$ 表示长度 l 在间隔的位置取到. 下面证明在所有情形下均有

$$|E(X_1X_{t_1}X_{t_2}X_{t_3})|\leqslant 8\alpha(\ell)^{\delta/(4+\delta)}\|X_1\|_{4+\delta}^4,$$

然后估计数 C_r.

$1\overset{\ell}{—}t_1t_2t_3$ 的情形：在命题 8.7.5 中，令 $p=4+\delta$，$q=(4+\delta)/3$，则 $1-1/p-1/q=\delta/(4+\delta)$，从而有

$$|E(X_1X_{t_1}X_{t_2}X_{t_3})-E(X_1)E(X_{t_1}X_{t_2}X_{t_3})|$$
$$\leqslant 8\alpha(\ell)^{\delta/(4+\delta)}\|X_{t_1}\|_{4+\delta}\|X_1X_{t_2}X_{t_3}\|_{(4+\delta)/3}$$
$$\leqslant 8\alpha(\ell)^{\delta/(4+\delta)}\|X_1\|_{4+\delta}^4,$$

其中最后一个估计式$\|X_1X_{t_2}X_{t_3}\|_{(4+\delta)/3}\leqslant\|X_1\|_{4+\delta}\|X_{t_2}\|_{4+\delta}\|X_{t_3}\|_{4+\delta}$由推广的 Hölder 不等式得到，见附录 A.4.1(ⅳ).

$1t_1t_2\overset{\ell}{—}t$ 的情形：与上一情形完全类似.

311

$1t_1\overset{\ell}{—}t_2t_3$ 的情形：取 $p=(4+\delta)/2$，$q=(4+\delta)/2$，从而 $1-1/p-1/q=\delta/(4+\delta)$，由此可得

$$|E(X_1X_{t_1}X_{t_2}X_{t_3})-E(X_1X_{t_1})E(X_{t_2}X_{t_3})|$$
$$\leqslant 8\alpha(\ell)^{\delta/(4+\delta)}\|X_1X_{t_1}\|_{(4+\delta)/2}\|X_{t_2}X_{t_3}\|_{(4+\delta)/2}$$
$$\leqslant 8\alpha(\ell)^{\delta/(4+\delta)}\|X_1\|_{4+\delta}^4,$$

其中，再一次用到了推广的 Hölder 不等式.

为了估计 C_r，由条件 $t_h-t_{h-1}=\ell$ 得 t_h 等于 $t_{h-1}+\ell$(ℓ 是固定的). 因为 t_i-t_{i-1} 在 $i=h$ 取到最大值 ℓ，其他每个指标 t_i，$i\neq h$ 只取到 $\ell+1$ 个值 t_{i-1}，$\cdots t_{i-1}+\ell$，所以 $C_r=\sum_{h=1}^{3}\sum_{n,3}^{(h)}1\leqslant n(\ell+1)$. 因此，可以得到

$$\begin{aligned}E(S_n^4)&=4!n\sum_{h=1}^{3}\sum_{n,3}^{(h)}|E(X_1X_{t_1}X_{t_2}X_{t_3})|\\&\leqslant 4!8n\|X_1\|_{4+\delta}\sum_{\ell=0}^{n-1}C_r\alpha(\ell)^{\delta/(4+\delta)}\\&\leqslant 4!8n\|X_1\|_{4+\delta}\sum_{\ell=0}^{\infty}(r+1)\alpha(\ell)^{\delta/(4+\delta)}\\&\leqslant 4!8A_4(\delta)\|X_1\|_{4+\delta}^4.\end{aligned}$$

■

评注 8.7.8 如果存在某个 $\delta>0$，使得 $A_r(\delta)<\infty$，则定理 8.7.7 的结论：$E|S_n|^r=O(n^{r/2})$ 对实值 $r>2$ 也成立，见 Yokoyama(1980).

最后，引入下列引理，它表明混合系数也可成为条件期望的界.

引理 8.7.9 设时间序列 $\{X_t: t\in\mathbf{Z}\}$ 满足 $E(X_0)=0$，且存在某个 $\delta>0$，使得 $E|X_0|^{2+\delta}<\infty$，记 $\mathcal{F}_{-n}=\sigma(X_t: t\leqslant -n)$，则

(ⅰ) $$\|E(X_0|\mathcal{F}_{-n})\|_{L_2}\leqslant 8\|X_0\|_{2+\delta}\alpha(n)^{\frac{\delta}{2(2+\delta)}},$$

(ⅱ) 如果

$$\sum_{n=1}^{\infty}\|E(X_0\mid\mathcal{F}_{-n})\|_{L_2}<\infty,$$

则有

312

$$\sum_{n=1}^{\infty}\alpha(n)^{\frac{\delta}{2(2+\delta)}}<\infty.$$

证明 注意到，对 $p=2$，$q=2+\delta$，有 $1/p-1/q=\delta/2(2+\delta)$. 因此，对任一满足 $\|Y\|_{L_2}<\infty$ 的 $\mathcal{F}_{-n}$ 可测的 Y，有

$$|E(X_0Y)|\leqslant 8\|X_0\|_{2+\delta}\|Y\|_{L_2}\alpha(n)^{\frac{\delta}{2(2+\delta)}}.$$

由引理 8.1.6 知

$$\begin{aligned}\|E(X_0\mid\mathcal{F}_{-n})]\|_{L_2}&=\sup\{E(X_0Y): Y\text{ 是}\mathcal{F}_{-n}\text{ 可测的}, \|Y\|_{L_2}=1\}\\&\leqslant 8\|X_0\|_{2+\delta}\alpha(n)^{\frac{\delta}{2(2+\delta)}},\end{aligned}$$

这就证明了(ⅰ). 由

$$\sum_{n=1}^{\infty}\|E(X_0\mid\mathcal{F}_{-n})\|_{L_2}\leqslant 8\|X_0\|_{L_2}\sum_{n=1}^{\infty}\alpha(n)^{\frac{\delta}{2(2+\delta)}}.$$

立即得到(ⅱ)成立. ■

8.8 混合过程的极限定理

本节讨论 α 混合时间序列的强大数定律.

定理 8.8.1(α 混合过程的强大数定律) 设$\{X_t\}$是 4 阶弱平稳过程，且是 α 混合的，其混合系数满足

$$A_4(\delta) = \sum_{i=0}^{\infty}(i+1)\alpha(i)^{\delta/(4+\delta)} < \infty,$$

对某个 $\delta>0$. 则当 $T\to\infty$时，

$$\frac{1}{T}\sum_{t=1}^{T}X_t \xrightarrow{\text{a. s.}} E(X_1).$$

证明 不失一般性，设 $E(X_1)=0$. 令 $\overline{X}_T = \frac{1}{T}\sum_{t=1}^{T}X_t, S_T = \sum_{t=1}^{T}X_t$. 考虑事件

$$A = \{\omega \in \Omega : \overline{X}_T(\omega) \to 0, T \to \infty\}.$$

下面证明 $P(A)=1 \Leftrightarrow P(A^C)=0$. 注意到

$$\begin{aligned}\overline{X}_T(w)\ \text{不收敛} \Leftrightarrow &\quad \exists\,\varepsilon > 0 : \forall\, T_0 \in \mathbf{N} : \exists\, T \geqslant T_0 : |\overline{X}_T(\omega)| > \varepsilon, \\ \Leftrightarrow &\quad A_T = \{|\overline{X}_T| > \varepsilon\}\ \text{无穷多次发生}.\end{aligned}$$

由 Borel-Cantelli 引理，如果

$$\sum_{T=1}^{\infty}P(A_T) < \infty$$

成立，则

$$P(A_T\ \text{无穷多次发生}) = P\Big(\bigcap_{T}\bigcup_{S\geqslant T}A_S\Big) = 0.$$

313

由推广的 Markov 不等式可得

$$P(A_T) = P(|\overline{X}_T| > \varepsilon) \leqslant \frac{ES_T^4}{T^4\varepsilon^4}.$$

根据定理 8.7.7(ii)，右边是 $O(T^{-2})$，由此可得 $\sum_{t=1}^{T}P(A_T) < \infty$. ■

定理 8.8.2(混合过程的中心极限定理) 设$\{X_t\}$是一个严平稳的 α 混合过程，满足

$$E(X_1) = 0 \quad 且 \quad E|X_1|^{2+\delta} < \infty$$

对某个 $\delta>0$. 若

$$\sum_{k=1}^{\infty}\alpha(k)^{\frac{\delta}{2+\delta}} < \infty,$$

则当 $T\to\infty$时，

$$\frac{1}{\sqrt{T}}\sum_{t=1}^{T}X_t \xrightarrow{d} N(0,\sigma^2),$$

其中

$$\sigma^2 = E(X_1) + 2\sum_{h=1}^{\infty} E(X_1 X_{1+h})$$

绝对收敛.

证明 因为$\{X_t\}$是严平稳的 L_2 过程，因此满足遍历性定理. 由引理 8.7.9，由关于混合系数的条件可得

$$\sum_{n=1}^{\infty} \| E(X_0 | \mathcal{F}_{-n}) \|_2 < \infty.$$

因此，可以直接应用定理 8.1.7. ■

下面建立下列样本自协方差函数的中心极限定理：

314

$$\hat{\gamma}_T(h) = \frac{1}{T}\sum_{t=1}^{T-h}(X_{t+h} - \overline{X}_T)(X_t - \overline{X}_T),$$

$0 \leqslant h \leqslant T-1$，否则$\hat{\gamma}_T(h)=0$；且

$$\widetilde{\gamma}_T(h) = \frac{1}{T-h}\sum_{t=1}^{T-h}(X_{t+h} - \mu)(X_t - \mu),$$

$0 \leqslant h \leqslant T-1$，否则$\widetilde{\gamma}_T(h)=0$.

定理 8.8.3 设$\{X_t\}$是严平稳的 α 混合过程，满足 $E(X_1)=\mu$ 且自协方差函数 $\gamma=\gamma_X$. 假定存在某个 $\delta>0$ 使得下列两个条件成立：

(ⅰ) $\sum_{k=1}^{\infty}\alpha(k)^{\frac{\delta}{2+\delta}} < \infty$.

(ⅱ) $E|X_1|^{4+4\delta} < \infty$.

那么，对任一固定的 h，当 $T\to\infty$ 时有

$$\sqrt{T}(\widetilde{\gamma}_T(h) - \gamma(h)) \xrightarrow{d} N(0, \eta_h^2),$$

其中

$$\eta_h^2 = \lim_{T\to\infty} E\left(\frac{1}{\sqrt{T}}\sum_{t=1}^{T} X_t X_{t+h}\right)^2.$$

证明 不失一般性，设 $\mu=0$. 固定 $h\geqslant 0$，考虑过程

$$\xi_t = X_t X_{t+h}, \quad t \in \mathbf{N}.$$

则 $E(\xi_t)=\gamma(h)$对所有 t 成立. 显然，$\{\xi_t\}$是严平稳的 L_2 过程，这是因为

$$\begin{aligned} E|\xi_t|^{2+\delta} &= E|X_t X_{t+h}|^{2+\delta} \\ &\leqslant \sqrt{E|X_t|^{2(2+\delta)}}\sqrt{E|X_{t+h}|^{2(2+\delta)}} \\ &= E|X_t|^{4+4\delta} < \infty. \end{aligned}$$

记$\{\xi_t\}$的第 k 个 α 混合系数为 $\alpha_\xi(k)$. 注意到

$$\mathcal{F}_{t,0} = \sigma(X_s X_{s+h} : s \leqslant t) \subset \sigma(X_s : s \leqslant t+h)$$

且

$$\mathcal{F}_{t,k} = \sigma(X_s X_{s+h} : s \geqslant t+k) \subset \sigma(X_s : s \geqslant t+k).$$

所以，对 $k>h$，可以由下式估计 $\alpha_\xi(k)$

$$\alpha_{\xi}(k)=\sup_{A\in\mathcal{F}_{t,0},B\in\mathcal{F}_{t,k}}|P(A\cap B)-P(A)P(B)|$$
$$\leqslant\sup_{A\in\sigma(X_s:s\leqslant t+h),B\in\sigma(X_s:s\geqslant t+k)}|P(A\cap B)-P(A)P(B)|$$
$$=\alpha(k-h).$$

315

由此可得

$$\sum_{k=1}^{\infty}\alpha_{\xi}(k)^{\frac{\delta}{2+\delta}}=\sum_{k=1}^{h}\alpha_{\xi}(k)^{\frac{\delta}{2+\delta}}+\sum_{k=h+1}^{\infty}\alpha_{\xi}(k)^{\frac{\delta}{2+\delta}}$$
$$=\sum_{k=1}^{h}\alpha_{\xi}(k)^{\frac{\delta}{2+\delta}}+\sum_{k=1}^{\infty}\alpha(k)^{\frac{\delta}{2+\delta}}<\infty.$$

最后，由定理 8.8.2 立即得到所需结论. ■

下列命题表明$\hat{\gamma}_T(h)$保留了$\widetilde{\gamma}_T(h)$的渐近分布性质.

命题 8.8.4 设$\{X_t\}$是一个弱平稳序列，且满足中心极限定理：当 $T\to\infty$ 时

$$\frac{1}{\sqrt{T}}\sum_{t=1}^{T}X_t\xrightarrow{d}N(0,\sigma^2),$$

则当 $T\to\infty$ 时

$$\sqrt{T}(\hat{\gamma}_T(h)-\widetilde{\gamma}_T(h))\xrightarrow{P}0.$$

证明 由直接计算可得

$$\hat{\gamma}_T(h)=\frac{1}{T}\sum_{t=1}^{T}(X_t-\overline{X}_T)(X_{t+h}-\overline{X}_T)$$
$$=\frac{1}{T}\Big(\sum_{t=1}^{T-h}X_tX_{t+h}-2\overline{X}_T\sum_{t=1}^{T-h}X_{t+h}+\overline{X}_T^2\Big).$$

因此，

$$\hat{\gamma}_T(h)-\widetilde{\gamma}_T(h)=-\overline{X}_T^2+2\overline{X}_T\frac{1}{T}\sum_{t=T-h+1}^{T}X_t-\frac{h}{T}\overline{X}_T^2,$$

从而得到

$$\sqrt{T}(\widetilde{\gamma}_T(h)-\widetilde{\gamma}_T(h))=-[\sqrt{T}\,\overline{X}_T]\,\overline{X}_T+2[\sqrt{T}\,\overline{X}_T]\frac{1}{T}\sum_{t=T+h+1}^{T}X_t-[\sqrt{T}\,\overline{X}_T]\frac{h}{T}\overline{X}_T.$$

由假设，括号内的项是渐近正态的，所以是 $O_P(1)$的. 剩下的部分是 $o_P(1)$的，这是因为由中心极限定理知，当 $T\to\infty$时，$\overline{X}_T\to 0$(依概率). ■

作为命题 8.8.4 的一个结论，当$\widetilde{\gamma}_T(h)$被$\hat{\gamma}_T(h)$代替时，即时间序列被样本均值 $\overline{X}_T$ 中心化时，定理 8.8.3 仍然成立. 进一步，人们可以把结论推广到依分布联合收敛的情形. 假设$\{X_t\}$是一个线性过程，有独立同分布的扰动项，且系数$\{\theta_i\}$满足定理 8.6.3 的条件，则有

316

$$\sqrt{T}[(\hat{\rho}_T(1),\cdots,\hat{\rho}_T(d))'-(\rho(1),\cdots,\rho(d))']\xrightarrow{d}N(0,W),$$

其中 $W=(w_{ij})_{i,j}$是$(d\times d)$-矩阵，其元素项为

$$w_{ij}=\sum_{k=-\infty}^{\infty}[\rho(k+i)\rho(k+j)+\rho(k-i)\rho(k+j)+2\rho(i)\rho(j)\rho^2(k)$$
$$-2\rho(i)\rho(k)\rho(k+j)-2\rho(j)\rho(k)\rho(k+i)].$$

特别地，对于一个白噪声过程，可以得到 $W=\mathrm{diag}(1,\cdots,1)$，这就验证了 3.3.1 节讨论的渐近假设检验和置信区间.

值得注意的是 $4+\delta$ 阶有限矩的假设是必要的. 事实上，如果是重尾分布的，那么样本自协方差和自相关可能有完全不同的性质.

有了自协方差的一致估计量，用它们来估计缩放样本均值的(渐近)方差是自然的，

$$\Sigma_T = \mathrm{Var}\Big(\frac{1}{\sqrt{T}}\sum_{t=1}^{T} X_t\Big) = \Gamma_0 + \sum_{j=1}^{T-1}\frac{T-j}{T}(\Gamma_j + \Gamma_j'),$$

其中$\{X_t: t=1,2\cdots\}$是一个平稳的 d 维时间序列，其均值为 μ，自协方差矩阵为

$$\Gamma_j = E(X_1-\mu)(X_{1+j}-\mu)', \quad j=1,2,\cdots.$$

对单变量情形情形进行推广，见(8.1)式，称

$$\Sigma = \lim_{T\to\infty}\Sigma_T = \Gamma_0 + \sum_{j=1}^{\infty}(\Gamma_j+\Gamma_j')$$

为**长期方差**(如果它存在).

在应用中，需要估计 Σ 的 d 维时间序列$\{X_t\}$通常表示为

$$X_t = h(Z_t;\vartheta),$$

其中$\{Z_t\}$是另一个严平稳的 l 维时间序列，$\vartheta\in\Theta$ 是参数，$h(z,\vartheta)$是定义在 $\mathbf{R}^l\times\Theta$ 上的已知函数，满足

$$E|h(Z_1;\vartheta)|^{4+\delta}<\infty,$$

对某个 $\delta>0$.

为了简化记号，令

$$h_t(\vartheta) = h(Z_t;\vartheta), \quad t\in\mathbf{N},$$

且假定存在某个真参数 $\vartheta_0\in\Theta$，使得

$$Eh_t(\vartheta_0) = 0, \quad \text{对所有的 } t$$

317 成立.

Σ 的 **Newey-West 估计量**定义为

$$\hat{\Sigma}_T = \hat{\Gamma}_{T0} + \sum_{j=1}^{m} w(j,m)(\hat{\Gamma}_{Tj}+\hat{\Gamma}_{Tj}'), \tag{8.28}$$

其中 $0\leqslant m<T$ 是**滞后截断参数**且

$$\hat{\Gamma}_{Tj} = \frac{1}{T}\sum_{t=j+1}^{T}\hat{h}_t(\vartheta_0)\hat{h}_{t-j}(\vartheta_0)'.$$

为了表述的简便，下面忽略记号中的参数，因为固定 $\vartheta=\vartheta_0$. 滞后截断参数选取为样本容量 $m=m_T$ 的函数，下面将会看到，它的阶必须小于 T.

上述的估计量依赖于随机变量 h_t 的估计. 考虑下列两种情形：

(ⅰ) 若 $h_t(\vartheta)=h(Z_t)$不依赖于 ϑ，则简单地使用$\hat{h}_t=h(Z_t)$.

(ⅱ) 若 $h_t(\vartheta)=h(Z_t,\vartheta)$，则假定已经有了某个一致估计量$\hat{\vartheta}_T$，且使用

$$\hat{h}_t = h(Z_t;\hat{\vartheta}_T), \quad t=1,2,\cdots.$$

注意，估计量$\hat{\Sigma}_T$ 只用到$\hat{h}_1,\cdots,\hat{h}_T$ 的样本自协方差矩阵的前 m 个对角元，它们被用

Bartlett 权重

$$w(j,m)=1-\frac{j}{m+1},\quad j=0,\cdots,m$$

进行加权. 这些权重满足下列正则性条件:

(ⅰ) $w(j,\ m)\leqslant C_w$, 对所有的 $0\leqslant j\leqslant m$, $m\in\mathbf{N}$, 及某个常数 $0<C_w<\infty$.

(ⅱ) 对每个 $0\leqslant j\leqslant m$, 当 $m\to\infty$ 时, $w(j,\ m)\to 1$.

如果令 $h_T=m+1$ 并引入 Bartlett 核函数

$$k_B(x)=(1-|x|)\mathbf{1}(|x|\leqslant 1),\quad x\in\mathbf{R}.$$

则有

$$w(j,m)=k_B\left(\frac{j}{h_T}\right),\quad j=0,\cdots,m,m\geqslant 1,$$

也可以选择其他核函数来定义其他的加权方案. 下面列出一些最常见的核函数.

(ⅰ) **截断核**: $k_{TR}(x)=\mathbf{1}(|x|\leqslant 1),x\in\mathbf{R}.$ 318

(ⅱ) **Parzen 核**:

$$k_P(x)=\begin{cases}1-6x^2+6|x|^3, & 0\leqslant|x|\leqslant 1/2,\\ 2(1-|x|)^3, & 1/2\leqslant|x|\leqslant 1,\\ 0, & \text{其他情形}.\end{cases}$$

(ⅲ) **Tukey-Hanning 核**:

$$k_{TH}(x)=\begin{cases}(1+\cos(\pi x))/2, & |x|\leqslant 1,\\ 0, & \text{其他情形}.\end{cases}$$

(ⅳ) **二次谱核**:

$$k_{QS}(x)=\frac{25}{12\pi^2x^2}\left(\frac{\sin(6\pi x/5)}{6\pi x/5}-\cos(6\pi x/5)\right).$$

为了得到估计 $\hat{h}_t$ 的假定, 设 $h(z;\ \vartheta)$ 是二阶可微的, 且满足

$$\sup_t\|\nabla h(Z_t;\vartheta_0)\|=O_P(1),\quad \sup_t\|Dh(Z_t;\vartheta_0)\|=O_P(1).$$

由泰勒公式可得

$$\hat{h}_t-h_t=\nabla h(Z_t;\vartheta_0)(\hat{\vartheta}_T-\vartheta_0)+O_P(\|\hat{\vartheta}_T-\vartheta_0\|).$$

如果 $\hat{\vartheta}_T$ 是 $\sqrt{T}$ 一致的, 即有

$$\sqrt{T}(\hat{\vartheta}_T-\vartheta_0)=O_P(1),$$

则估计 $\hat{h}_t$ 是依速率 $\sqrt{T}$ 的均匀一致估计, 即

$$\max_{t\leqslant T}\sqrt{T}|\hat{h}_t-h_t|=O_P(1).$$

为了证明 Newey-West 估计量的一致性, 它需要下列极大不等式.

命题 8.8.5(α 混合序列的极大不等式) 设 $\{\xi_t:\ t\in\mathbf{N}\}$ 是一个零均值的 α 混合时间序列, 存在某个 $\delta>0$ 使得

$$E|\xi_t|^{2+\delta}<\infty$$

对所有的 $t\geqslant 1$ 成立, 且混合系数 $\{\alpha(k):\ k\in\mathbf{N}\}$ 满足

319

$$\sum_{k=1}^{\infty}\alpha(k)^{\frac{2}{2+\delta}}<\infty.$$

那么

$$P\left(\max_{1\leqslant t\leqslant T}\left|\sum_{t=1}^{T}\xi_t\right|>\varepsilon\right)\leqslant\frac{4\sum_{t=1}^{T}E\xi_t^2+c(\alpha,\delta)\sum_{t=1}^{T}(E|\xi_t|^{2+\delta})^{\frac{\delta}{2+\delta}}}{\varepsilon^2},$$

其中

$$c(\alpha,\delta)=16\left[(4\delta^{-1}+2)\sum_{k=1}^{\infty}(k+1)^{\frac{2}{\delta}}\alpha(k)^{\frac{\delta}{2+\delta}}\right].$$ ■

定理 8.8.6(Newey-West 估计量的一致性) 设$\{Z_t\}$是严平稳的α混合时间序列，其混合系数满足

$$\sum_{k=1}^{\infty}\alpha(k)^{\frac{\delta}{2+\delta}}<\infty.$$

进一步，设$h_t(\vartheta)=h(Z_t;\ \vartheta)$，$t\geqslant1$，其中，函数$h(z;\ \vartheta)$满足

$$Eh(Z_1;\vartheta_0)=0,\quad E|h(Z_1;\vartheta_0)|^{4+4\delta}<\infty.$$

如果$\hat{h}_t$是h_t的均匀一致估计，即

$$\max_{t\leqslant T}|\hat{h}_t-h_t|=o_P(1).\tag{8.29}$$

那么，如果滞后截断法则满足当$T\to\infty$时，

$$\frac{m^2}{T}=o(1),$$

那么Σ_T的 Newey-West 估计量$\hat{\Sigma}_T$是一致的，即当$T\to\infty$时，

$$\hat{\Sigma}_T-\Sigma_T\xrightarrow{P}0.$$

证明 为了简化记号，设$h_t=h_t(\vartheta_0)$. 注意到$\hat{\Sigma}_T$可以写为

$$\hat{\Sigma}_T=\frac{1}{T}\sum_{t=I}^{T}\hat{h}_t\hat{h}_t'+\frac{2}{T}\sum_{j=1}^{m}\sum_{t=j+1}^{T}w(j,m)\hat{h}_t\hat{h}_{t-j}'.$$

下面先证

320

$$\check{\Sigma}_T=\frac{1}{T}\sum_{t=1}^{T}h_th_t'+\frac{2}{T}\sum_{j=1}^{m}\sum_{t=j+1}^{T}w(j,m)h_th_{t-j}'$$

是一致的，由于$\hat{\Sigma}_T$是$\check{\Sigma}_T$中用随机变量h_t的估计量$\hat{h}_t$代替h_t，然后证明用h_t的估计量代替h_t后仍是一致的. 考虑与Σ_T对应的截断形式

$$\widetilde{\Sigma}_T=\frac{1}{T}\sum_{t=1}^{T}E(h_th_t')+\frac{2}{T}\sum_{j=1}^{m}\sum_{t=j+1}^{T}w(j,m)E(h_th_{t-j}'),$$

注意到

$$\widetilde{\Sigma}_T=E(\check{\Sigma}_T).$$

如能证明下列结论：

(ⅰ) $\|\Sigma_T-\widetilde{\Sigma}_T\|=o(1)$.

(ⅱ) $\|\check{\Sigma}_T-\widetilde{\Sigma}_T\|=O(m^2/T)$.

(ⅲ) $\|\hat{\Sigma}_T-\check{\Sigma}_T\|=o_P(1)$.

则定理得证. 下面记$\|\cdot\|$为矩阵的极大值范数. 先验证(i)，即估计误差$\|\Sigma_T-\widetilde{\Sigma}_T\|$. 因为有

$$\|\Sigma_T-\widetilde{\Sigma}_T\|\leqslant\frac{2}{T}\sum_{j=1}^{m}\sum_{t=j+1}^{T}|w(j,m)-1|\,\|E(h_th'_{t-j})\|+\frac{2}{T}\sum_{j=m+1}^{T}\sum_{t=j+1}^{T}\|E(h_th'_{t-j})\|.$$

由 α 混合条件知第二项满足：当 $m\to\infty$时

$$2C\max_{N\leqslant T}\sum_{j=1}^{N}\frac{N-j}{N}\alpha(j)^{\frac{2}{2+\delta}}=O\Big(\sum_{j=m+1}^{\infty}\alpha(j)^{\frac{2}{2+\delta}}\Big)=o(1).$$

这是因为 $\sum_{j=1}^{\infty}\alpha(j)^{\frac{2}{2+\delta}}<\infty$，且由 $Eh_t=0$ 和命题 8.7.6 知随机矩阵 $h_th'_{t-j}$的(k,l)元素 $h_{tk}h_{t-j,l}$满足

$$|E(h_{tk}h_{t-j,l})|\leqslant C\alpha(j)^{\frac{2}{2+\delta}},$$

对某个不依赖于 k，$l\in\{1,\cdots d\}$的常数 $0<C<\infty$. 用类似的方法，第一项的界为

$$B_m=c\sum_{j=1}^{m}|w(j,m)-1|\alpha(j)^{\frac{2}{2+\delta}},$$

对某个常数 $0<c<\infty$. 定义函数序列

$$f_n(j)=|w(j,n)-1|\alpha(j)^{\frac{2}{2+\delta}}\mathbf{1}(j\leqslant n),\quad j\in\mathbf{N}_0,$$

$n\in\mathbf{N}$，则

$$B_m=c\int f_m(x)\mathrm{d}v(x),$$

321

其中 $v(x)$表示 $\mathbf{N}_0$ 上的计数测度. 现在由控制收敛定理可得 $B_m=o(1)$，因为对每个固定的 $x\in\mathbf{N}$ 有 $f_m(x)\to0$，$m\to\infty$，且

$$f_m(x)\leqslant g(x)=2\alpha(x)^{\frac{2}{2+\delta}},$$

其中

$$\int g(x)\mathrm{d}v(x)=2\sum_{k=1}^{\infty}\alpha(x)^{\frac{2}{2+\delta}}<\infty.$$

为了证明(ii)，注意到

$$\check{\Sigma}_T-\widetilde{\Sigma}_T=R_{T1}+R_{T2},$$

其中

$$R_{T1}=\frac{1}{T}\sum_{t=1}^{T}(h_th'_t-E(h_th'_t)),$$

$$R_{T2}=\frac{2}{T}\sum_{j=1}^{m}w(j,m)\sum_{t=j+1}^{T}[h_th'_{t-j}-E(h_th'_{t-j})]$$

$$\leqslant\frac{2C_w}{T}\sum_{j=1}^{m}\Big|\sum_{t=j+1}^{T}[h_th'_{t-j}-E(h_th'_{t-j})]\Big|.$$

由命题 8.8.5 给出的极大不等式，可得 $R_{T^2}=O(m^2/T)$. 现固定 $1\leqslant j\leqslant T-1$，考虑序列

$$U_{tj}=h_th'_{t-j}-E(h_th'_{t-j}),\quad t\geqslant1.$$

记 U_{tj}的第(k,l)元素为 $U_{tj}^{(k,l)}$，$1\leqslant k$，$l\leqslant d$，由命题 8.8.5 及联合界知，可以按下列式子去估计 R_{T^2}，

$$
\begin{aligned}
P(|R_{T2}|>\varepsilon) &\leqslant \sum_{k,l=1}^{d}\sum_{j=1}^{m}P\Big(\max_{1\leqslant N\leqslant T}\frac{1}{N}\Big|\sum_{t=j+1}^{N}U_t^{(k,l)}\Big|>\frac{\varepsilon}{C_w m}\Big)\\
&\leqslant \sum_{k,l=1}^{d}\sum_{j=1}^{m}C_w^2 m^2\,\frac{4TE|U_l^{(k,l)}|^{4+2\delta}+c(\alpha,\delta)T(E|U_1^{(k,l)}|^{4+2\delta})^{\frac{2}{2+\delta}}}{\varepsilon^2T^2}\\
&=O\Big(\frac{m^2}{T}\Big).
\end{aligned}
$$

从而用估计$\widetilde{h}_t$代替h_t后仍是一致的，因为(8.29)式表明

322

$$
\hat{h}_t=h_t+\frac{\sqrt{T}(\hat{h}_t-h_t)}{\sqrt{T}}=h_t+O_P(1/\sqrt{T}),
$$

在$t\leqslant T$时是一致，由此得到

$$
\begin{aligned}
\hat{h}_t\,\hat{h}_{t-j}&=[h_t+O_P(1/\sqrt{T})][h'_{t-j}+O_P(1/\sqrt{T})]\\
&=h_th'_{t-j}+O_P(1/\sqrt{T}),
\end{aligned}
$$

其中$O_P(1/\sqrt{T})$项在$j\leqslant T-1$和$t\leqslant T$时是一致的，因为由假设h_t的四次绝对矩是一致有界的. 所以

$$
\frac{1}{T}\sum_{t=1}^{T}\hat{h}_t\,\hat{h}_{t-j}=\frac{1}{T}\sum_{t=1}^{T}h_th'_{t-j}+O_P(1/\sqrt{T}),
$$

这就验证了(ⅲ)，证毕. ■

8.9 评注与延伸阅读

对侧重于参数估计的时间序列渐近分布理论感兴趣的读者可阅读 Brockwell 和 Davis (1991). 专著(Davidson，1994，p. 383)给出了定理 8.3.6(鞅差分序列的中心极限定理)的一个生动而巧妙的证明(本书未给出)，它是通过验证一个 McLeish(1974)的基本定理的条件来证明的. 关于混合过程的内容，本书参考了 Durrett(1996)，Bosq(1998)和 Doukhan (1994). 值得一提的是定理 8.1.7 中的条件 $\sum_{n=1}^{\infty}\|E(X_0|\mathcal{F}_{-n})\|_2<\infty$ 可以减弱为 $\sum_{n=1}^{\infty}\|E(X_0|\mathcal{F}_0)-E(X_0|\mathcal{F}_{-1})\|_2<\infty$，见(Hall 和 Heyde，1980，定理 5.3)以及 Durrett (1996). 定理 8.6.3(线性过程的中心极限定理)有一个更弱条件 $\sum_i|\theta_i|<\infty$，见(Hannan，1970，定理 11)，在 Merlevede 等人的在评论文章中，改进了证明的方法. 关于半鞅极限定理的高等概率论专著，读书可以参考 Jacod 和 Shiryaev(2003). 线性过程的渐近性理论可以在 Phillips 和 Solo(1992)中找到. Davidson(2000)从计量经济学的角度探讨了带随机回归量的多元线性回归模型的渐近理论. 对于非参数密度估计，非参数回归及其应用，我们推荐 Silverman(1986)，Hardl(1990)，Fan 和 Gijbels(1996)，Fan 和 Yao(2003)，Li 和 Racine(2007)以及 Franke(2008)等著作. 要对本章内容有更深入的了解，可阅读计量经济学专著 White(2001)和 Tanaka(1996). 定理 8.8.5(极大不等式)选自(Peligrad，1999，推论 2.4). Bartlett 估计几乎必然收敛的一般结论可参考 Berkes 等(2005).

参考文献

Berkes I., Horváth L., Kokoszka P. and Shao Q.M. (2005) Almost sure convergence of the Bartlett estimator. *Period. Math. Hungar.* **51**(1), 11–25.

Bosq D. (1998) *Nonparametric Statistics for Stochastic Processes*. vol. 110 of *Lecture Notes in Statistics* 2nd edn. Springer-Verlag, New York. Estimation and prediction.

Brockwell P.J. and Davis R.A. (1991) *Time Series: Theory and Methods*. Springer Series in Statistics 2nd edn. Springer-Verlag, New York.

Davidson J. (1994) *Stochastic Limit Theory: An Introduction for Econometricians*. Advanced Texts in Econometrics. The Clarendon Press Oxford University Press, New York.

Davidson J. (2000) *Econometric Theory*. Blackwell Publishing.

Doukhan P. (1994) *Mixing*. vol. 85 of *Lecture Notes in Statistics: Properties and Examples*. Springer-Verlag, New York.

Durrett R. (1996) *Probability: Theory and Examples*. 2nd edn. Duxbury Press, Belmont, CA.

Fan J. and Gijbels I. (1996) *Local Polynomial Modelling and its Applications*. vol. 66 of *Monographs on Statistics and Applied Probability*. Chapman & Hall, London.

Fan J. and Yao Q. (2003) *Nonlinear Time Series: Nonparametric and Parametric Methods*. Springer Series in Statistics. Springer-Verlag, New York.

Franke J., Härdle W.K. and Hafner C.M. (2008) *Statistics of Financial Markets: An Introduction*. Universitext 2nd edn. Springer-Verlag, Berlin.

Hall P. and Heyde C.C. (1980) *Martingale Limit Theory and its Application*. Academic Press Inc. [Harcourt Brace Jovanovich Publishers], New York. Probability and Mathematical Statistics.

Hannan E.J. (1970) *Multiple Time Series*. John Wiley and Sons, Inc., New York-London-Sydney.

Härdle W. (1990) *Applied Nonparametric Regression*. vol. 19 of *Econometric Society Monographs*. Cambridge University Press, Cambridge.

Jacod J. and Shiryaev A.N. (2003) *Limit Theorems for Stochastic Processes*. vol. 288 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]* 2nd edn. Springer-Verlag, Berlin.

Li Q. and Racine J.S. (2007) *Nonparametric Econometrics: Theory and Practice*. Princeton University Press, Princeton, NJ.

McLeish D.L. (1974) Dependent central limit theorems and invariance principles. *Ann. Probability* **2**, 620–628.

Merlevède F., Peligrad M. and Utev S. (2006) Recent advances in invariance principles for stationary sequences. *Probab. Surv.* **3**, 1–36 (electronic).

Peligrad M. (1999) Convergence of stopped sums of weakly dependent random variables. *Electron. J. Probab.* **4**, no. 13, 13 pp. (electronic).

Phillips P.C.B. and Solo V. (1992) Asymptotics for linear processes. *Ann. Statist.* **20**(2), 971–1001.

Silverman B.W. (1986) *Density Estimation for Statistics and Data Analysis: Monographs on Statistics and Applied Probability*. Chapman & Hall, London.

Tanaka K. (1996) *Time Series Analysis*. Wiley Series in Probability and Statistics: Probability and Statistics. John Wiley & Sons Inc., New York. Nonstationary and noninvertible distribution theory, A Wiley-Interscience Publication.

White H. (2001) *Asymptotic Theory for Econometricians*. Academic Press, San Diego, CA.

Yokoyama R. (1980) Moment bounds for stationary mixing sequences. *Z. Wahrsch. Verw. Gebiete* **52**(1), 45–57.

323 ~ 324

第9章 几个专题

本章介绍几个专题. 首先，讨论用 copula 函数(连接函数)方法对高维多元分布进行分析及建模，这在金融中已经被广泛使用了，其中的一个重要应用是对担保债务凭证(CDO)进行定价，借此机会本章讨论了 copula 如何演变成 2008 年金融危机的帮凶.

非参数方法已经成为金融数据分析一个不可分割的部分. 在很多领域，局部多项式法，特别是局部线性估计量，在简单与准确之间提供了很好的平衡. 因此，本章将详细地介绍局部多项式法所需的渐近理论，它是在前面几章讨论的基础上进行的补充和延伸.

最后，本章选择性地讨论几种变点方法. 这些方法假设利率时间序列可能存在结构性断点(变点)，此时它们的分布也随之改变，这种情况是相当符合实际的，因为大量的潜在因素可能对价格、收益、风险测度、指数或其他决定利率的量产生影响，如盈利预警、合并、融合、政治事件、极端天气等这种意料之外的新闻可能导致分布的改变. 在这种情况下，就要求即时行动，比如投资组合更新、对冲、风险出售或平仓. 用于检验相应变化时刻的时间序列分析以及应用监控程序来快速检测变点的变点检验理论，已经成为当前的研究热点.

9.1 copula 和 2008 年的金融危机

copula 已经成为对随机向量的多元分布(例如联合信用违约时间)进行建模的常选方法，copula 使得人们能够将问题分成若干特定的问题或估计它们的边际分布及处理坐标间的相关性. 很多 copula 模型只有几个参数甚至是一个参数的模型. 由此，一些棘手的问题(比如有几百或几千个金融工具的复杂风险测度的计算)可以化为相当简单的公式和程序，这种方法在处理债务时显得特别重要. 违约事件通常是相关的，特别是当市场上出现危机时，一个违约事件常常导致下一个违约. 通常，没有足够的历史数据来分别估计每个信贷和债务人的违约概率. copula 模型使人们能够通过几个参数来刻画相关性，这些参数有时与有效的市场信息有关.

325

为这种简约付出的代价是：对一个给定的 copula 函数，其相关性结构是固定的，一维模型对一个比如 1000 维分布来说只代表了所有可能分布的一个微小子集，如果被选取的 copula 模型是错误，并且不符合实际，则由模型计算得到的风险测度和价格将极具误导性. 这正是发生在 2008 年金融危机之前和期间的事情. 它们被越来越多地用于担保债务凭证和信用违约互换的定价. Li 模型成为行业标准并且被评级机构和银行大量使用. Li 模型依赖于一个简单的高斯 copula 模型，在处理问题时非常简单，但用这一模型计算得出的违约概率过于乐观了.

在引入 copula 及其基本性质后，本章将分析导致这场金融风暴的主要原因，然后详细介绍著名的 Li 模型.

9.1.1 copula

设 $X=(X_1, \cdots, X_d)$是一个 d 维随机向量，其中 $d\in \mathbf{N}$ 可以非常大，它的多元分布函数为

$$F(x_1,\cdots,x_d)=P(X_1\leqslant x_1,\cdots,X_d\leqslant x_d),\quad x_1,\cdots,x_d\in\mathbf{R}.$$

copula 方法的基本思想如下：设 $C(u_1,\cdots,u_d)$ 是随机向量 $(U_1,\cdots,U_d)$ 的分布函数，其中，$(U_1,\cdots,U_d)$ 有相同的均匀边际分布，即 $U_i\sim U(0,1)$，$i=1,\cdots,d$，则

$$C(F_1(x_1),\cdots,F_d(x_d)),\quad x_1,\cdots,x_d\in\mathbf{R}\tag{9.1}$$

定义了一个 $\mathbf{R}^d$ 上的分布函数，其边际分别为 $F_1,\cdots,F_d$. 事实上，有 $\lim\limits_{x\to\infty}F_j(x)=1$，$j=1,\cdots,d$，且

$$\lim_{u_2,\cdots,u_d\to\infty}C(u_1,\cdots,u_d)=C(u_1,1,\cdots,1)=u_1,\quad u_1\in[0,1],$$

从而得到

$$\lim_{x_2,\cdots,x_d\to\infty}C(F_1(x_1),\cdots,F_d(x_d))=F_1(x_1),\quad x_1\in\mathbf{R}.$$

这就证明了第一个坐标成立，其他坐标类似可证.

定义 9.1.1(copula) 一个 **copula** 是一个 d 维随机向量 $(U_1,\cdots,U_d)$ 的分布函数 $C:[0,1]^d\to[0,1]$，其中 $U_i\sim U(0,1)$，$i=1,\cdots,d$.

326

设 $X\sim F$ 是一个随机变量，对 $x\in\mathbf{R}$，$a>0$ 定义

$$F(x,a)=P(X\leqslant x)+aP(X=x).$$

设 V 为一个随机变量，且满足 $V\sim U(0,1)$，则称

$$U=F(X,V)=F(X-)+V[F(X)-F(X-)]$$

为一个**分布变化或概率积分变化**. 注意，对一个连续分布函数 F，有 $F(x,a)=F(x)$，在这种情况下分布变化可简单地由 $U=F(x)$ 给出. 分布变化满足

$$U\sim U(0,1),\quad X=F^{-1}(U)(\text{a. s.}).$$

下面验证后一事实，由定义

$$F(X-)\leqslant U=F(X,V)\leqslant F(X)$$

且 $P(U=F(X-))=P(V=0)$. 因为 $F^{-1}(u)=x$ 对所有的 $u\in(F(x-),F(x)]$ 成立，所以 $F^{-1}(U)=X$ (a. s.).

下面的著名 Sklar 定理表明 $\mathbf{R}^d$ 上的任意分布函数可以表示成式(9.1)的形式，它是 copula 方法的基础.

定理 9.1.2(Sklar 定理，1959)

(i) 设 F：$\mathbf{R}^d\to[0,1]$ 是一个 d 维分布函数，且 $F_i(i=1,\cdots,d)$ 是对应的边际分布函数. 则存在一个 d 维分布函数 C：$[0,1]^d\to[0,1]$ 满足

$$F(x_1,\cdots,x_d)=C(F_1(x_1),\cdots,F_d(x_d))$$

对所有的 $(x_1,\cdots,x_d)'\in\mathbf{R}^d$ 成立.

(ii) 如果 C：$[0,1]^d\to[0,1]$ 是一个 copula 函数，并且 $F_1,\cdots,F_d$ 是单变量分布函数，则

$$C(F_1(x_1),\cdots,F_d(x_d)),\quad (x_1,\cdots,x_d)\in\mathbf{R}^d$$

定义了一个 d 维分布函数，其边际分布函数为 $F_1,\cdots,F_d$.

证明 先证(i). 设 $(X_1,\cdots,X_d)$ 是一个随机向量，其分布函数为 F. 对一个独立于 X 的随机变量 $V\sim U(0,1)$，考虑分布变化 $U_i=F_i(X_i,V_i)$ 得到 $X_i=F_i^{-1}(U_i)(\text{a. s.})$，$i=$

1，…，d. copula 函数

$$C(u_1,\cdots,u_d)=P(U_1\leqslant u_1,\cdots,U_d\leqslant u_d),\quad u_1,\cdots,u_d\in[0,1]$$

满足

$$\begin{aligned}F(x_1,\cdots,x_d)&=P(X_1\leqslant x_1,\cdots,X_d\leqslant x_d)\\&=P(F_1(U_1)\leqslant x_1,\cdots,F_d(U_d)\leqslant x_d)\\&=P(U_1\leqslant F_1^{-1}(x_1),\cdots,U_d\leqslant F_d^{-1}(x_d))\end{aligned}$$

327 对所有的 x_1，…，$x_d\in\mathbf{R}$ 成立. ■

如果一个多元分布由 copula 函数 C 表示，其边际分布函数为 F_1，…，F_d，那么可以按下列方法模拟分布函数为 $F(x_1,\cdots,x_d)=C(F_1(x_1),\cdots,F_d(x_d))$的随机向量$(X_1,\cdots,X_d)$：

1)通过 copula 函数 C，选取一个随机向量$(U_1,\cdots,U_d)\sim C$.

2)由每个分量的分位数函数 F_i^{-1} 获取随机向量

$$(X_1,\cdots,X_d)=(F_1^{-1}(U_1),\cdots,F_d^{-1}(U_d)),$$

其分布函数为 F.

3)重复步骤 1 和 2，n 次后得到一组 d 维随机向量的随机样本 $X^{(1)}$，…，$X^{(n)}\sim F$.

注意到，如果 Sklar 定理中的 d 维分布函数 F 有连续的边际分布函数 F_1，…，F_d，则 copula 函数满足

$$C(u_1,\cdots,u_d)=F(F_1^{-1}(u_1),\cdots,F_d^{-1}(u_d))$$

对每个 $u=(u_1,\cdots,u_d)'\in\mathbf{R}^d$ 成立. 且对每个 $u=(u_1,\cdots,u_d)'\in\mathbf{R}^d$，有

$$C(u_1,\cdots,u_d)=P(F_1(X_1)\leqslant u_1,\cdots,F_d(X_d)\leqslant u_d).$$

下面列举一些常用的 copula 函数.

例 9.1.3(copula 函数)

(ⅰ) **独立性**：设 U_1，…，$U_d\sim U(0,1)$是独立的，则对应的自变量为 u_1，…，u_d 的独立 copula 函数为

$$P(U_1\leqslant u_1,\cdots,U_d\leqslant u_d)=\prod_{i=1}^d P(U_i\leqslant u_i)=\prod_{i=1}^d u_i.$$

(ⅱ) **完全相关**：若 U_1，…，$U_d\sim U(0,1)$是完全相关的，$U_1=\cdots=U_d$，则对应的 copula 函数为

$$C_u(u_1,\cdots,u_d)=P(U_1\leqslant\min(u_1,\cdots,u_d))=\min(u_1,\cdots,u_d),$$

u_1，…，$u_d\in[0,1]$，称它为**上 Fréchet copula 函数**.

(ⅲ) **混合 copula 函数**：设 $C_0(u_1,\cdots,u_d)$为任一 copula 函数(比如，独立 copula 函数)且 $\rho>0$ 是一个混合系数，组合 $C_0(u_1,\cdots,u_d)$与完全相关对应的 copula 函数得到一个新的 copula 函数，

$$C(u_1,\cdots,u_d)=(1-\rho)C_0(u_1,\cdots,u_d)+\rho\min(u_1,\cdots,u_d),$$

328 u_1，…，$u_d\in[0,1]$.

(ⅳ) **二元高斯 copula 函数**：对 $d=2$，考虑二维正态分布，每个边际分布的均值为 0，方差为 1，它们之间的相关系数 $\rho\in(-1,1)$，则 copula 函数为

$$C(u_1,u_2)=\int_{-\infty}^{\Phi^{-1}(u_1)}\int_{-\infty}^{\Phi^{-1}(u_2)}\frac{1}{2\pi(1-\rho^2)^{1/2}}\Big(-\frac{s^2-2\rho st+t^2}{2(1-\rho^2)}\Big)\mathrm{d}s\mathrm{d}t,$$

u_1，$u_2\in[0, 1]$. 注意，这是带有一个参数的函数族.

(ⅴ) **二元 t-copula 函数**：

$$C(u_1,u_2)=\int_{-\infty}^{F_v^{-1}(u_1)}\int_{-\infty}^{F_v^{-1}(u_2)}\frac{1}{2\pi(1-\rho^2)^{1/2}}\left(1+\frac{s^2-2\rho st+t^2}{v(1-\rho^2)}\right)^{-(v+2)/2}\mathrm{d}s\mathrm{d}t,$$

u_1，$u_2\in[0, 1]$，其中 $F_v^{-1}(p)$是自由度为 v 的 t 分布的分位数函数，$\rho\in(-1, 1)$是相关系数.

(ⅵ) **高斯 copula 函数**：二元高斯 copula 函数可以容易地推广到 d 维. 一般地，称

$$C(u_1,\cdots,u_d)=\Phi_\Sigma(\Phi^{-1}(u_1),\cdots,\Phi^{-1}(u_d)),\quad u_1,\cdots,u_d\in[0,1]$$

为一个高斯 copula 函数，其中 Φ_Σ 是均值为 0，协方差阵为 Σ 的一个多元正态分布的分布函数，Φ^{-1}是 $N(0, 1)$分布的分位数函数. 对 $d=2$，$\Sigma=\begin{pmatrix}1 & \rho\\ \rho & 1\end{pmatrix}$，则得到二元高斯 copula 函数.

(ⅶ) **极值 copula 函数**：如二维分布对应的 copula 函数为

$$C(u_1,u_2)=\exp\left[\log(xy)A\left(\frac{\log(u_1)}{\log(u_1u_2)}\right)\right],\quad u_1,u_2\in[0,1],$$

则称这种 copula 函数为极值 copula 函数，其中 A：$[0, 1]\to[1/2, 1]$是一个凸函数，且 $A(t)$满足条件：$\max(t, 1-t)\leqslant A(t)\leqslant 1$ 对所有的 $t\in[0, 1]$成立.

(ⅷ) **Kimedorf-Sampson copula 函数**：称

$$C(u_1,u_2)=(u_1^{-\gamma}+u_2^{-\gamma}-1)^{1/\gamma},\quad u_1,u_2\in[0,1]$$

$(0\leqslant\gamma<\infty)$为 Kimedorf-Sampson copula 函数.

(ⅸ) **Archimedean copula 函数**：称一个 copula 函数为阿基米德(Archimedean) copula 函数，如它有如下形式

$$C(u_1,\cdots,u_d)=\psi(\psi^{-1}(u_1)+\cdots+\psi^{-1}(u_d)),$$ 329

这里称函数 ψ 为**生成元**. 为了确保 C 是一个 copula 函数，生成元必须在$[0, \infty)$上是 d 阶单调的，即

$$(-1)^k\psi(k)(x)\geqslant 0$$

对所有的 $x\geqslant 0(k=0, 1, \cdots, d-2)$成立，且$(-1)^{d-2}\psi^{(d-2)}(x)$是非增的凸函数. 常用的生成元及它们的逆函数如下.

Ali-Mikhail-Haq 生成元：生成元为

$$\psi(t;\theta)=\frac{1-\theta}{\exp(t)-\theta},\quad \theta\in[0,1),$$

其逆函数为

$$\psi^{-1}(t;\theta)=\log\frac{1-\theta+\theta t}{t}.$$

Clayton 生成元：生成元为

$$\psi(t;\theta)=(1+t)^{-1/\theta},\quad \theta\in(0,\infty),$$

其逆函数为

$$\psi^{-1}(t;\theta)=t^{-\theta}-1.$$ 330

Frank 生成元：生成元为

$$\psi(t;\theta)=-\frac{\log(1-(1-\exp(-\theta))\exp(-t))}{\theta},\quad \theta\in(0,\infty),$$

其逆函数为

$$\psi^{-1}(t;\theta)=-\log\frac{\exp(-\theta t)-1}{\exp(-\theta)-1}.$$

■

图 9-1 和图 9-2 表明 DAX 和 FTSE 的每日对数收益(其边际由一个核密度估计量进行非参数估计得到)的最大似然估计的高斯 copula 函数得到一个基于 copula 函数的二维分布估计量.

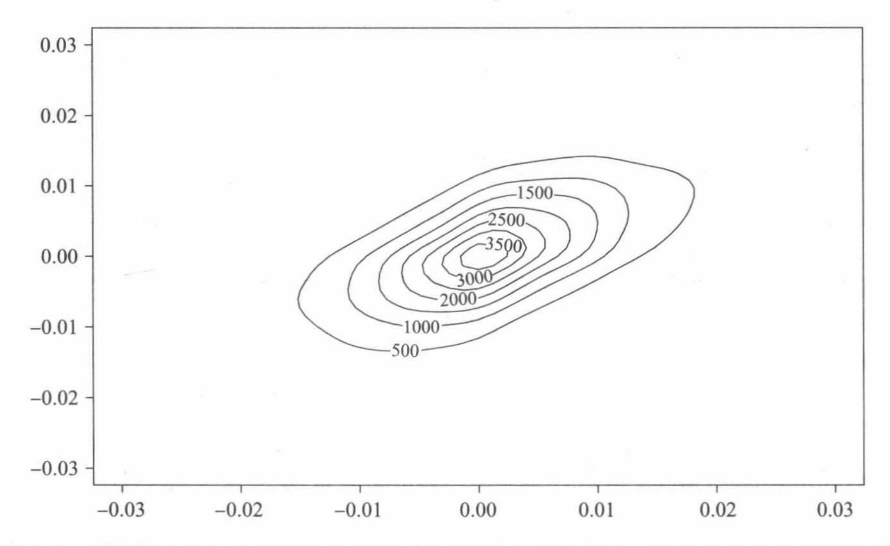


图 9-1 用高斯函数和边际核估计得到的 DAX 和 FTSE 每日对数收益分布的等值线图

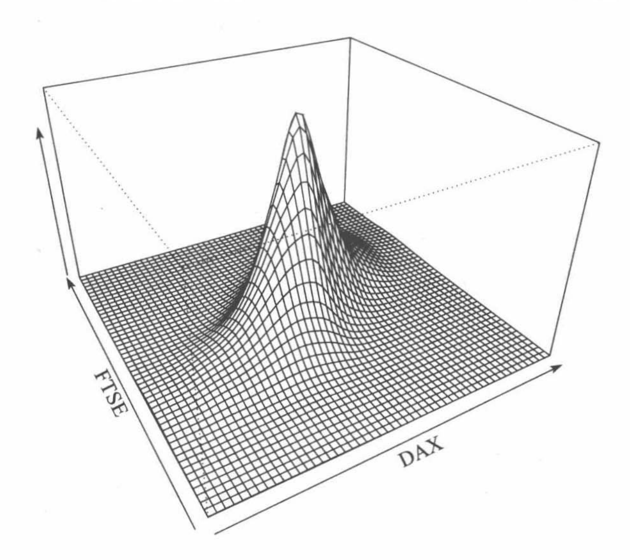


图 9-2 由高斯 copula 函数和边际核估计构建的 DAX 和 FTSE 每日对数收益分布的 3D 图像

一个分布函数为 $F(x_1, \cdots, x_d)=C(F_1(x_1), \cdots, F_d(x_d))$的 d 维随机向量上的样本 $X_1, \cdots, X_T$

$$X_t = (X_{t1},\cdots,X_{td})', \quad t = 1,\cdots,T,$$

对应的 copula 函数 $C(u_1, \cdots, u_d)$的非参数估计可由下列步骤得到. 为了简化叙述，设边际分布函数 $F_1, \cdots, F_d$ 是连续的. 因为

$$C(u_1,\cdots,u_d) = P(F_1(X_1) \leqslant u_1,\cdots,F_d(X_d) \leqslant u_d),$$

$u_1, \cdots, u_d \in [0, 1]$. 如果知道了变化 $U_{ij}=F_j(X_{ij})$，则 copula 函数一个自然的估计量是

$$V_T(u_1,\cdots,u_d) = \frac{1}{T}\sum_{t=1}^{T}\mathbf{1}(U_{t1} \leqslant u_1,\cdots,U_{td} \leqslant u_d), \quad u_1,\cdots,u_d \in [0,1].$$

331

由于 F_j 是未知的，用它们的非参数估计量

$$\hat{F}_{Tj}(x) = \frac{1}{T}\sum_{t=1}^{T}\mathbf{1}(X_{tj} \leqslant x), \quad x \in \mathbf{R}$$

代替它们，得到 $U_{tj}=F_j(X_{tj})$的估计量为

$$\hat{U}_{tj} = \hat{F}_{Tj}(X_{tj}), \quad j = 1,\cdots,d;t = 1,\cdots,T.$$

注意到 $T\hat{U}_{tj}$是所有满足 $X_{t'j}<X_j$ 的 $X_{t'j}(t'=1, \cdots, T)$的和，即 X_{tj}的第 j 个坐标的样本 $X_{1j}, \cdots, X_{Tj}$的秩 R_{tj}，这些秩几乎必然是数 1，…，T 的一个排列. 现未知的 copula 函数的非参数估计为

$$\hat{C}_T(u_1,\cdots,u_d) = \frac{1}{T}\sum_{i=1}^{T}\mathbf{1}(\hat{F}_{T1}(X_{i1}) \leqslant u_1,\cdots,\hat{F}_{Td}(X_{id}) \leqslant u_d), \quad u_1,\cdots,u_d \in [0,1].$$

由于 $\hat{F}_{T1}(X_{i1})\leqslant u_1 \Leftrightarrow X_{i1}\leqslant \hat{F}_{T1}^{-1}(u_1)$，这一估计量也可以写为

$$\hat{C}_T(u_1,\cdots,u_d) = \frac{1}{T}\sum_{i=1}^{n} \mathbf{1}(X_{i1} \leqslant \hat{F}_{Td}^{-1}(u_1),\cdots,X_{id} \leqslant \hat{F}_{T1}^{-1}(u_d)).$$

当 $T\to\infty$时，对应的经验过程

$$\sqrt{T}[\hat{C}_T(u_1,\cdots,u_d) - C(u_1,\cdots,u_d)], \quad u_1,\cdots,u_d \in [0,1]$$

弱收敛于高斯过程

$$G_C(u_1,\cdots,u_d) = B_C(u_1,\cdots,u_d) - \sum_{j=1}^{d} \frac{\partial C}{\partial u_j}(u_1,\cdots,u_d)B_C(u_1,\cdots,u_{j-1},1,u_{j+1},\cdots,u_d),$$

$u_1, \cdots, u_d\in[0, 1]$，其中 B_C 是$[0, 1]^d$ 上的一个布朗桥，其协方差函数为 $\mathrm{Cov}(B_C(u_1, \cdots, u_d)), B_C(u_1', \cdots, u_d'))$等于

$$C(\min(u_1,u_1'),\cdots,\min(u_d,u_d')) - C(u_1,\cdots,u_d)C(u_1',\cdots u_d')$$

其中 $u_1, \cdots, u_d, u_1', \cdots, u_d'\in[0, 1]$.

9.1.2 金融危机

从 20 世纪 90 年代中期到 2006 年，全美房价每年上涨，因此形成了一个巨大的泡沫. 在那段时间，低利率带动了抵押贷款的增多和房价的增高，这也刺激了金融机构建立一些

金融工具来提高收益．从2000年开始，由于次级抵押贷款的快速增加，泡沫逐渐膨胀了．
332 可调整利率的抵押贷款(ARM)被发明出来用于低收入、信用不够好的人，否则，这些人将被抵押市场和投机者排除在外．ARM通常在最初几年提供优惠的利率并且不需要首付，比如为首付提供第二个按揭．有时，借款人还被允许推迟偿还部分到期的利息并将其添加到本金．这些交易背后的逻辑是预期房价会持续上涨，这样就可以允许借款人(投机者)在几年后通过利率优惠用一个新按揭来再融资，或者以更高的价格卖出房产．但是美国房价上涨幅度从2005年4月开始下降，这使得早期的预期开始落空．很多次贷借款人的抵押贷款成本大幅提高，结果，至2006年，次贷(ARM)的拖欠率上升到10.09%，而同期优质固定利率抵押贷款的拖欠率仅为2.27%．

次级抵押贷款的增加被一种新的金融创新推波助澜了，这种金融创新将抵押转化为一种标准化的金融工具，称为*证券化过程*．两家被评为类似AAA债券并大量从事按揭贷款业务的政府资助企业，Fannie Mae和Freddie Mac，在20世纪70年代开发了*房产抵押贷款证券*(MBS)，这种证券将各种抵押打包并增加担保使它们能吸引人．通过这种方式，Fannie Mae和Freddie Mac能够从全国各地银行得到贷款并将相应的风险卖给风险偏好投资者，从而各地银行对客户扩大按揭贷款．MBS吸引人的思想是按揭根据不同的风险状况被重新打包成不同的份额，在这种情况下，按揭市场空前繁荣．由于全国范围内的违约事件从未发生过，MBS就被认为是一种相当安全的投资．虽然两家企业均被监管，并且规定它们只能从事具有一定信用等级的借款人的普通贷款，但是私营企业开发了次贷抵押来支持MBS，它比标准抵押有更高的收益．

如何构建MBS是值得关注的．MBS交易包含被现金流联系的两方：资产方是投资组合的基础方，负债方是证券的发行者，通常是一个特殊的机构(SPV)．SPV是由资金所有者创造的一种具有法人资格的实体，目的是将投资者与信用风险的源头(通常是一家银行)分隔开．银行将资产卖给一个SPV，然后SPV发行一种票据．票据根据损失的百分比被分为高级份额、中层份额、中低级份额和股权份额．比如百分比为3%，4%，5%，85%．股权份额须承担最初3%的损失，中低级份额承担4%的损失，以此类推．比如由于一个违约债务，当投资组合有5%的损失时，股权份额须承担3%的损失，中低级份额承担2%．其他两个份额不受影响．这样，高级份额持有人承担的风险比股权份额持有人低．进一步，对于按揭产生的收益，高级份额具有优先权，它们最先被支付．一旦高级份额持有人得到支付，中级份额持有人接下来得到支付，如此下去，剩余的收益才分配给股权份额．这样，级别越高的股权份额就越安全．MBS进行了*超额抵押信用增级*，也就是说抵押资产的面值高于再打包证券的面值，而超额抵押部分就是股权份额．在上面的例子中，按揭付
333 款超过3%违约，更高份额才会有损失．

份额的再打包并没有减小基础投资组合的风险，它只是进行了重新组合．高级份额具有高投资级别的信贷评级，因为它们差不多与违约风险隔绝了．但是低等级的份额却具有大得多的风险，并且会很快遭受损失．

2000年左右，商业银行和投资银行创建了一种新的金融工具来证券化次级抵押贷款，它极大地加速了泡沫的扩大：抵押及其他资产支持的*担保债务凭证*(CDO)打包证券．CDO发行

者购买不同的 MBS 份额并将它们与其他资产支持的证券比如信用卡贷款、汽车贷款、学生贷款汇集在一起. 所以一个 MBS 的资产是由真实的按揭付款组成，而一个 CDO 资产却是集合这些按揭付款的证券. 显然，投资者所承受的风险变得不透明了. CDO 在 2006 年和 2007 年被大量地发行. 但没有足够的 ABS 可以交易，大量的 CDO 支持了合成的 ABS. 在金融危机前，发行者直接同评级机构一起合作开发 CDO 份额并购买信用违约互换或信用保险来使得他们的 CDO 获得高评级. 实际上，这些 CDO 的高级份额被错误定价了：投资银行通过发行只需支付较低收益的 AAA 级 CDO 债券支付较低的收益来购买高收益的房产抵押贷款证券(MBS). 通过简单地再打包现金流，就有可能产生一个正的净现值. 但是当 2007 年 CDO 的降级波在大范围冲击市场时，由于 CDO 是被过于简单的 Li 模型错误定价的，先前高等级的份额面临巨大的损失，从而 CDO 的持有者面临困境了. 投资者比如只能用一定量的资金投资于 AAA 以下债券的养老基金，就有以任何价格抛售其持有的压力.

这曾经是一种积极有效的反馈机制：在泡沫期间当价格持续上涨时，头寸按市值计价的银行净值也随之增加，这就增加了它们的杠杆. 因此，银行着重于运用其剩余资本，即扩大资产负债表，因为当价格上涨时保持低杠杆的银行是无利可图的. 这就进一步加剧了对按揭类相关产品的需求. 次级资产的供应顺应了这种持续增长的需求. 当危机发生时，杠杆快速放大，缺乏流动性的金融机构在账面上有大量包含坏账的证券，而这些资产再也无法流通交易了.

银行持有部分 MBS、CDO 以及其他结构化投资工具(SIV)债券. 银行创造了 SPV 并持有这些资产来增加杠杆头寸，这比它们受监管机构资本要求的资产负债表的头寸要多. 为了融资到 SIV 头寸，SPV 以负债形式发行资产支持商业票据(ABCP). 这些工具绝大部分都是小于 2 星期的短期限并且必须持续拆借. 然而，当危机爆发时，很多银行再一次把 SIV 放到资产负债表上.

隔夜回购协议(回购贷款)变成投资银行短期借款的一种流行工具. 银行用资产作为抵押和另外一家银行进行隔夜贷款. 这样，金融机构就被联系在一起了，当一家银行陷入困境时，问题将蔓延到其他机构. 据估计，隔夜回购在总资产中的份额从 12%扩张到 25%. 在 2007 年雷曼兄弟 37%的负债是抵押借款，22%是短头寸的，这些短期工具原本是为长期资产融资的. 这些短期流动性资金的干涸(特别是 ABSP 和回购贷款)作为银行间不信任的结果，成为这次金融危机中的一个重要元素.

334

增强 MBS 和 CDO 的一个自然想法是信用保险. 在 20 世纪 70 年代，Ambac 这类高信用级别的单一险种保险公司开始支持市政债券发行. 通过提供违约保险，市政能够以 AAA 等级借款. 由于市政债券的违约概率被高估了，单一险种便利用了这一评级套利. 通过卖出信用违约互换为 CDO 和其他房产抵押贷款证券持有者保险，这一商业模式被扩大到住房市场.

当然，这类金融工具支持的大类信用投资组合的大量交易需要定价公式来确定它们的公平价值. David. X. Li 用高斯 copula 函数来刻画相关性并建议使用信用违约互换利差来校准在市场中的价值. 上面已经提到，这个模型变成金融行业的标准并且也被评级机构使用. 当金融危机爆发时，由于相关债权人的违约风险变大，CDS 迅速扩散增加，而前者反

过来又导致 CDO 份额价值的暴跌. 评级机构对 CDO 的评级开始适应现实的需要. 2007 年 6 月，穆迪对价值 50 亿美元住宅按揭贷款支持的次级 MBS 降级，并重审 184 CDO 份额. 标准普尔对价值 73 亿美元的 MBS 进行降级观察.

9.1.3 信用违约模型和 CDO

本节从一些初等微积分开始，设 T 是一个固定的时间，A 和 B 分别表示两个债务人或两家公司违约两笔贷款的事件，设 T 时刻前违约的概率分别为 $p_A=P(A)=E(\mathbf{1}_A)$ 和 $p_B=P(B)=E(\mathbf{1}_B)$，联合违约事件 $A\cap B$ 发生的概率为 $p_{AB}=P(A\cap B)$. 当违约是相互独立的时候，$p_{AB}=p_Ap_B$，否则两起违约就是相关的，因此计算其条件概率

$$p_{A|B}=\frac{p_{AB}}{p_B},\quad p_{B|A}=\frac{p_{AB}}{p_A}$$

及相关系数 $\rho_{AB}=\mathrm{Cor}(\mathbf{1}_A, \mathbf{1}_B)$ 是有意义的，显然有

$$\rho_{AB}=\frac{p_{AB}-p_Ap_B}{\sqrt{p_A(1-p_A)p_B(1-p_B)}}.$$

设 n 个债务人借了 n 笔贷款，面值分别为 N_1，…，N_n. 令 r_1，…，r_n 为回收率，即若贷款 i 违约，仍可以得到支付 r_iN_i，$i=1$，…，n，并假定回收率是非随机的.

令

$$\tau_1,\cdots,\tau_n:(\Omega,\mathcal{F},Q)\to[0,\infty)$$

分别为 n 个债务人(借款人)的违约时间，借款人可以是个人也可以是公司，Q 为一个固定的定价概率测度. 分别用 F 和 S 表示联合分布的分布函数和生存函数，即对 $(x_1,\cdots,x_n)\in\mathbf{R}^n$

$$F(x_1,\cdots,x_n)=Q(\tau_1\leqslant x_1,\cdots,\tau_n\leqslant x_n)$$

与

335

$$S(x_1,\cdots,x_n)=Q(\tau_1>x_1,\cdots,\tau_n>x_n).$$

由 Sklar 定理知，存在一个 copula 函数 C 使得

$$F(x_1,\cdots,x_n)=C(F_1(x_1),\cdots,F_n(x_n)),$$

其中，F_1，…，F_n 分别是违约时间 τ_1，…，τ_n 的边际分布函数. 反之，如果给定了一个 copula 函数和边际函数 F_1，…，F_n，则可得到一个联合违约分布模型.

由于 CDO 通常是建立在大量的信贷基础上的，如果有历史数据，则可以估计(联合)违约概率，但是，通常历史数据是缺乏的. 另一个问题是维数 n 可能非常大，这在实证中甚至会阻碍相关性的估计，因为存在着 $n(n-1)/2$ 个两两相关系数. 例如，当 $n=100$ 时，需要估计 4590 个相关系数.

在基于资产模型或结构化模型中，基础债务人的价值 $V_i(t)$ 被模型化了. 在 $V_i(t)<K_i$ 的第一时刻，违约发生，即

$$\tau_i=\inf(t\geqslant 0:V_i(t)<K_i),$$

其中 K_i 是一个转换障碍，$i=1$，…，n. 同样地，信用降级可以通过选择这样的障碍建模. 相关的违约就是值 $V_1(t)$，…，$V_n(t)$ 的一列相关结果.

在**基于强度的模型**中，设存在 d 个非负函数 λ_1，…，λ_d：$[0, \infty)\to[0, \infty)$，使得第 i 个债务人(或贷款人)在 t 时刻存活的生存概率为

$$p_i(t) = P(\tau_i > t) = \exp\left(-\int_0^t \lambda_i(u)\mathrm{d}u\right), \quad t \in [0,\infty).$$

注意到 $p_i(t)$是非增的，对应的分布函数为

$$F_i(t) = 1 - p_i(t) = 1 - \exp\left(-\int_0^t \lambda_i(u)\mathrm{d}u\right), \quad t \in [0,\infty).$$

则违约时间为

$$\tau_i = \inf(t \geqslant 0 : p_i(t) \leqslant U_i), \tag{9.2}$$

其中随机变量 U_1，…，U_n 是单位区间$[0, 1]$上的均匀分布．事实上，因为对任意数 $p\in[0, 1]$和任意随机变量 $U\sim U(0, 1)$，有 $P(U<p)=p$．对(9.2)式中的 τ_i，因为$p_i(t)$是非增的，所以

$$\begin{aligned} P(\tau_i > t) &= P(p_i(s) > U_i, s \in [0,t]) \\ &= P(p_i(t) > U_i) \\ &= p_i(t), \end{aligned}$$

从而得到

$$P(\tau_i \leqslant t) = F_i(t), \quad t \in \mathbf{R}$$

对 $i=1$，…，d 成立．事实上，可以用 $F_i(\tau_i)(i=1, \cdots, d)$代替 U_i，因为这些随机变量都是$[0, 1]$上的均匀分布，违约间的相关性可以由 U_1，…，U_n 的相关性引出．Li 模型使用一个高斯 copula 函数，这意味着违约时间的联合分布函数由一个高斯 copula 函数 C_g 给出，即

336

$$P(\tau_1 \leqslant x_1, \cdots, \tau_d \leqslant x_d) = C_g(F_1(x_1), \cdots, F_d(x_d)), \quad x_1, \cdots, x_d \in \mathbf{R}. \tag{9.3}$$

设 d 家公司申请 d 项贷款或者发行 d 只投资组合的债券，其资产价值为 Z_i．单因子方法假设

$$V_i = \rho F + \sqrt{1-\rho^2}G_i, \tag{9.4}$$

其中 F，G_1，…，G_d 是在固定定价测度 Q 下独立的标准正态随机变量，F 是所有债务人共有的市场因素，且 $\rho\in[0, 1]$决定了受因素 F 的影响程度．若 $\rho=0$ 则违约是独立的，而 $\rho=1$ 则有同单调性质．随机变量 G_i 表示债务人的特殊风险．注意到若 $i\neq j$，$\mathrm{Cov}(V_i, V_j)=\rho$ 且 $\mathrm{Var}(V_i)=1$．因此，随机向量$(V_1, \cdots, V_n)$是多元正态分布，其均值为 0，协方差阵 $\Sigma(\rho)$的对角元为 1，非对角元为 ρ．这也就是说 ρ 是资产间的相关系数．

因为 $\tau_i\sim F$，所以 $F_i(\tau_i)\sim U(0, 1)$．在 Li 模型中，设标准正态随机变量 V_i 与违约时间 τ_i 的关系为

$$V_i = \Phi^{-1}(F_i(\tau_i)), \quad i = 1, \cdots, d. \tag{9.5}$$

注意到 $V_i=\Phi^{-1}(F_i(\tau_i))\Leftrightarrow F_i(\tau_i)=\Phi(V_i)$，所以有

$$\begin{aligned} P(\tau_1 \leqslant x_1, \cdots, \tau_d \leqslant x_d) &= P(F_1(\tau_1) \leqslant F_1(x_1), \cdots, F_d(\tau_d) \leqslant F_d(x_d)) \\ &= P(V_1 \leqslant \Phi^{-1}(F_1(x_1)), \cdots, V_d \leqslant \Phi^{-1}(F_d(x_d))) \\ &= C_{\Sigma(\rho)}(F_1(x_1), \cdots, F_d(x_d)). \end{aligned}$$

这意味着违约时间的联合分布由式(9.3)给出且 $C_g = C_{\Sigma(\rho)}$. 注意，上述讨论不需要式(9.5)成立，只需要假设

$$(V_1, \cdots, V_n) \overset{d}{=} (\Phi^{-1}(F_1(\tau_1)), \cdots, \Phi^{-1}(F_n(\tau_n))).$$

给定 F，时刻 t 的条件违约概率可以由下式计算.

$$\begin{aligned} p_i(t \mid V) &= Q(\tau_i \leqslant t \mid F) \\ &= Q(F_i^{-1}(\Phi(V_i)) \leqslant t \mid F) \\ &= Q(V_i \leqslant \Phi^{-1}(F_i(t)) \mid F). \end{aligned}$$

337

现在由模型(9.4)得到

$$\begin{aligned} p_i(t \mid V) &= Q(\rho F + \sqrt{1-\rho^2} G_i \leqslant \Phi^{-1}(F_i(t)) \mid V) \\ &= Q\Big(G_i \leqslant \frac{\Phi^{-1}(F_i(t)) - \rho F}{\sqrt{1-\rho^2}}\Big) \\ &= \Phi\Big(\frac{\Phi^{-1}(F_i(t)) - \rho F}{\sqrt{1-\rho^2}}\Big). \end{aligned}$$

为了用上述模型给 CDO 定价，需要估计相关系数 ρ，方法是标准化的. 信用违约互换模型中参数的选取是使得模型价格与市场价格相匹配.

高斯 copula 模型的简便使之在金融行业变得非常诱人，它迅速地被采用，并且成为了一个约定俗成的方法. 因为只有一个参数，它很容易被当做市场定价的标准，也使得大型投资组合的定价能快速地计算出来. 但它也有一些严重的缺陷. (联合)违约建模是不够的，违约往往成群地发生. 如果一家公司违约了，很可能另一家公司也违约，比如他们有相同的业务或者有相似的引起违约的外部风险.

更糟糕的是，在高斯 copula 函数下，当违约规模变大时，违约在上尾的相关性消失意义下是不相关的. 设 T_1 和 T_2 是两个违约时间，其边际分布函数分别为 F_1 和 F_2，如果极限

$$d = \lim_{p \to 1-} P(T_2 > F^{-1}(q) \mid T_1 > F^{-1}(q))$$

存在，则称该极限为**上尾相关系数**. 若 $d=0$，称 T_1 和 T_2 为**上尾渐近不相关的**. 系数 d 只依赖于 copula 函数. 对于相关性小于 1 的高斯 copula 函数，可以证明 $d=0$. 这意味着在高斯 copula 函数下，不管相关系数的值为多少，极端事件看起来都是不相关的.

9.2 局部线性非参数回归

下面讨论一般情形的非参数回归，它是已经学习过的内容的推广. 设 $\{(Y_t, X_t): t \in \mathbf{N}\}$ 是一个离散时间过程，它们满足

$$Y_t = m(X_t) + \varepsilon_t,$$
$$\varepsilon_t = \sigma(X_t)\xi_t,$$

其中 Y_t 是一元响应变量，X_t 是 d 维回归量，$d \in \mathbf{N}$，$\{\xi_t\}$ 是扰动项，满足

$$E(\xi_t \mid X_t) = 0, \quad E(\xi_t^2 \mid X_t) = 1,$$

对于函数 σ: $\mathbf{R}\to(0,\infty)$ 与 m: $\mathbf{R}\to\mathbf{R}$，且 $\sigma(x)$ 是连续的，$m(x)$ 存在 $p+1$ 阶偏导数，且 $p+1$ 阶偏导数有界，$p\in\mathbf{N}$. 引入滤子 338

$$\mathcal{F}_0=\{\varnothing,\Omega\},\quad \mathcal{F}_t=\sigma(Y_{s-1},X_s:s\leqslant t),\quad t\geqslant 0,$$

注意到

$$E(Y_t\,|\,X_t=x)=m(x)+\sigma(x)E(\xi_t\,|\,X_t=x)=m(x)$$

且

$$\operatorname{Var}(Y_t\,|\,X_t=x)=\sigma^2(x)E(\xi_t^2\,|\,X_t=x)=\sigma^2(x),\quad \text{a.s.},$$

这意味着 $m(x)$ 是回归函数，$\sigma(x)$ 是条件波动率.

9.2.1 金融中的应用：鞅测度估计和 Itô 扩散估计

上述方法可以解决金融中很多重要的估计问题：风险中性密度(等价鞅测度)的估计，自回归条件异方差非参数模型的估计和有离散样本的扩散过程的估计.

▶**例 9.2.1**(风险中性密度估计) 在第 1 章 1.5.8 节中，T 时刻股价 S_T 的风险中性密度 $\varphi_T^*(x)$ 与敲定价格为 K，到期日为 T 的欧式看涨期权的无套利价格 $C_e(K)$ 相关，关系为

$$\frac{\partial^2 C_e(K)}{\partial K^2}=\varphi_T^*(K).$$

设期权价格的一个样本为 C_1，…，C_n，到期日分别为 K_1，…，K_n，满足

$$C_i=m(K_i)+\varepsilon_i,\qquad i=1,\cdots,n,$$

其中，$m(K_i)=E(C_i\,|\,K_i)$ 且 $\varepsilon_i=C_i-m(K_i)$. 那么问题归结为用样本 (C_1, K_1)，…，(C_n, K_n) 估计二阶导数 $m^{(2)}(x)$. ◀

▶**例 9.2.2**(非参数 ARCH 模型) 记 $R_t=\log(P_t/P_{t-1})(t=1,\cdots,T)$ 为等间隔的 T 个时刻观察到的价格过程 $\{P_t\}$ 的对数收益. 设 $\{R_t\}$ 是一个均值为 0 的平稳时间序列，且满足一个 p 阶非参数 ARCH 模型，$p\in\mathbf{N}$

$$R_t=\sigma(R_{t-1}^2,\cdots,R_{t-p}^2)\xi_t,\quad t=1,2,\cdots,$$

其中 ξ_t 是随机变量，独立于

$$\mathcal{F}_{t-1}=\sigma(R_{t-1},\cdots,R_{t-p}),$$

且满足 $E(\xi_t)=0$ 及 $E(\xi_t^2)=1$，对所有的 t 成立，又设 $\sigma(x_1,\cdots,x_p)$ 是定义在 $[0,\infty)\times\cdots\times[0,\infty)$ 上的光滑函数. 注意到，$\sigma(R_{t-1},\cdots,R_{t-p})$ 是给定过去信息 $\mathcal{F}_{t-1}=\sigma(R_{t-1},R_{t-2},\cdots)$ 的条件波动率. 考虑 339

$$R_t^2=\sigma^2(R_{t-1},\cdots,R_{t-p})\xi_t^2,\quad t=1,2,\cdots.$$

则

$$E(R_t^2\,|\,\mathcal{F}_{t-1})=\sigma^2(R_{t-1},\cdots,R_{t-p}),$$

且可以用模型

$$R_t^2=\sigma^2(R_{t-1},\cdots,R_{t-p})+\varepsilon_t,\quad t=1,\cdots,T,$$

来估计 $\sigma^2(x_1,\cdots,x_p)$，其中 $\varepsilon_t=R_t^2-\sigma^2(R_{t-1},\cdots,R_{t-p})$ 的条件均值为 0. ◀

▶**例 9.2.3**(伊藤扩散的非参数估计) 设 $\{X(t): t\geqslant 0\}$ 是一个平稳遍历的伊藤扩散，即它是下列随机微分方程的一个平稳遍历解

$$dX(t) = \mu(X(t))dt + \sigma(X(t))dB_t,$$

其中 $X(0)$ 由平稳分布 μ 确定. 在 6.7 节中对这类过程是随机微分方程平稳遍历解的充分条件已有讨论，在那里还介绍了许多例子. 设时间步长为 $\Delta>0$，使用欧拉近似方法

$$X_t = X(t\Delta), \quad t = 1,2,\cdots,$$

由此得到方程

$$Y_t = \mu(X_{t-1})\Delta + \sigma(X_{t-1})\sqrt{\Delta}\varepsilon_t, \quad t = 1,2,\cdots,$$

其中

$$Y_t = X_t - X_{t-1}, \quad t = 1,2,\cdots$$

是一个平稳的时间序列. ◀

9.2.2 方法和渐近讨论

设 X_t 在 $\mathbf{R}$ 上取值，下面估计 $m(x)$ 在固定点 x 的值. 在 x 邻域内将 $m(x)$ 展成泰勒多项式，则有

$$m(z) \approx \sum_{k=0}^{p} \frac{m^{(k)}(x)}{k!}(z-x)^k.$$

固定某个单变量平滑核 K 和带宽 $h>0$. $m(x)$ 的**局部多项式估计量**及各阶导数 $m'(x)$，…，$m^{(p)}(x)$ 定义为

$$\hat{m}_T(x) = \hat{\delta}_{T0}(x),$$
$$\hat{m}_T^{(k)}(x) = k!\hat{\delta}_{Tk}(x), k = 1,\cdots,p,$$

340

其中 $\hat{\delta}_T(x) = (\hat{\delta}_{T0}，\cdots，\hat{\delta}_{Tp})'$ 为

$$\hat{\delta}_T(x) = \underset{(\delta_0,\cdots,\delta_p)\in\mathbf{R}^{p+1}}{\operatorname{argmin}} \sum_{t=1}^{T}\Big(Y_t - \sum_{j=0}^{p}\delta_j(X_t - x)^j\Big)^2 K\Big(\frac{X_t - x}{h}\Big).$$

定义

$$X = \begin{pmatrix} 1 & X_1 - x & (X_1 - x)^2 & \cdots & (X_1 - x)^p \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & X_T - x & (X_T - x)^2 & \cdots & (X_T - x)^p \end{pmatrix},$$
$$Y = (Y_1,\cdots,Y_T)',$$
$$W = \operatorname{diag}\Big(K\Big(\frac{X_1 - x}{h}\Big),\cdots,K\Big(\frac{X_T - x}{h}\Big)\Big).$$

则 $\hat{\delta}_T(x)$ 是下列线性方程的一个解

$$(X'WX)\hat{\delta}_T(x) = X'WY,$$

且有显式表达式

$$\hat{\delta}_T(x) = (X'WX)^{-1}X'WY.$$

注意到：$X'WX = (a_{ij})_{i,j}$ 是一个范德蒙德矩阵，其中元素为

$$a_{ij} = \sum_{t=1}^{T} K\Big(\frac{X_t - x}{h}\Big)(X_t - x)^{i+j}, 1 \leqslant i,j \leqslant p+1.$$

因为 X_t 是连续分布，所以矩阵是正规的(a. s.).

下面，将仅讨论 $p=1$ 时的局部线性估计量，但是允许多元情形，即 $d>1$. 为了定义平滑权，再一次使用乘积核 $K(u)=\prod_{j=1}^{d}L(u_j)$，$u=(u_1,\cdots,u_d)\in\mathbf{R}^d$，见(8.11)式，其中 L 是单变量平滑核，满足假定(8.9)和(8.10).

设回归函数 m：$\mathbf{R}^d\to\mathbf{R}$ 属于 C_b^3 类，考虑泰勒展式

$$m(X_t)=m(x)+\nabla m(x)(X_t-x)+\frac{1}{2}(X_t-x)'Dm(x)(X_t-x)+o(\|X_t-x\|^2),$$

其中$\nabla m(x)$为梯度，$Dm(x)$为二阶偏导数矩阵. 观察到

$$m(X_t)=\binom{1}{X_t-x}'\vartheta(x)+\frac{1}{2}(X_t-x)'Dm(x)(X_t-x)+o(\|X_t-x\|^2),$$

其中

$$\vartheta(x)=\binom{m(x)}{\nabla m(x)}.$$

341

$\vartheta(x)$的**局部线性估计量**为

$$\hat{\vartheta}_T(x)=\underset{\vartheta(x)\in\mathbf{R}^2}{\operatorname{argmin}}\sum_{t=1}^{T}K\left(\frac{X_t-x}{h}\right)\left(Y_t-\binom{1}{X_t-x}'\vartheta(x)\right)^2,$$

得到显示表达式

$$\hat{\vartheta}_T(x)=\left[\frac{1}{Th}\sum_{t=1}^{T}K\left(\frac{X_t-x}{h}\right)\binom{1}{X_t-x}(1,X_t-x)\right]^{-1}$$

$$\times\frac{1}{Th}\sum_{t=1}^{T}K\left(\frac{X_t-x}{h}\right)\binom{1}{X_t-x}Y_t,$$

这表明括号中的矩阵如果乘以一个对角元素为 1，h_1^{-2}，$\cdots$，h_d^{-2} 的对角矩阵 $Q_h(x)$，则它依概率收敛于一个正规矩阵. 因为对任意的数 w_t 和向量 a_t 有 $Q_h\sum_{t=1}^{T}w_ta_ta'_t=\sum_{t=1}^{T}w_ta_t(Q_ha_t)'$. 所以有

$$\hat{\vartheta}_T(x)=A_T^{-1}(x)\cdot\frac{1}{Th}\sum_{t=1}^{T}K\left(\frac{X_t-x}{h}\right)\begin{bmatrix}1\\ \dfrac{X_t-x}{h^2}\end{bmatrix}Y_t, \tag{9.6}$$

$$A_T(x)=\frac{1}{Th}\sum_{t=1}^{T}K\left(\frac{X_t-x}{h}\right)\begin{bmatrix}1\\ \dfrac{X_t-x}{h^2}\end{bmatrix}(1,X_t-x) \tag{9.7}$$

$$=\frac{1}{Th}\sum_{t=1}^{T}K\left(\frac{X_t-x}{h}\right)\begin{bmatrix}1 & (X_t-x)'\\ \dfrac{X_t-x}{h^2} & \dfrac{(X_t-x)(X_t-x)'}{h^2}\end{bmatrix}, \tag{9.8}$$

这是渐近分析的出发点. 在这里及以后，都假定 h^2 满足

$$\frac{X_t-x}{h^2}=\left(\frac{X_{ti}-x_i}{h_i{}^2}\right)_{i=1}^{d}.$$

这表明截距的收敛速度与斜率的收敛速度不同. 事实上,

(ⅰ) 对 $\delta_1(x)=m(x)$, $\hat{\vartheta}_{T1}(x)$ 是 $\sqrt{Th_1\cdots h_d}$ 一致的;

(ⅱ) $\hat{\vartheta}_{T2}(x)$ 是 $\sqrt{Th_1\cdots h_d}\operatorname{diag}(h_1, \cdots, h_d)$ 一致的.

这里 $\hat{\vartheta}_T=(\hat{\vartheta}_{T1}, \hat{\vartheta}'_{T2})'$, 其中 $\hat{\vartheta}_{T1}$ 估计截距, 而 $\hat{\vartheta}_{T2}$ 估计斜率. 右边的缩放矩阵是下列对角矩阵

342

$$H_T=\operatorname{diag}(\sqrt{Th_1\cdots h_d},\sqrt{Th_1\cdots h_d}h_1,\cdots,\sqrt{Th_1\cdots h_d}h_d).$$

下面叙述并证明局部线性估计量的渐近正态结果, 该定理在相当弱的条件下得到了平滑平均的中心极限定理.

定理 9.2.4(局部线性估计量的中心极限定理) 设 $\{(Y_t, X_t): t=1, 2, \cdots\}$ 是一个平稳遍历的 L_2 时间序列, 使得随机向量

$$\begin{pmatrix}1\\ \dfrac{X_t-x}{h^2}\end{pmatrix}\varepsilon_t,\quad t=1,2,\cdots,$$

满足下列中心极限定理

$$U_T(x)=\frac{1}{\sqrt{Th_1\cdots h_d}}\sum_{t=1}^{T}K\left(\frac{X_t-x}{h}\right)\begin{pmatrix}1\\ \dfrac{X_t-x}{h}\end{pmatrix}\varepsilon_t\xrightarrow{d}N(0,V),\quad T\to\infty,\tag{9.9}$$

其中 $\varepsilon_t=Y_t-E(Y_t|X_t)$, V 是某个矩阵, 且

$$\max_j h_j=o(1),\quad Th_1\cdots h_d\sum_{j=1}^{d}h_j^4=o(1).$$

又设 f 的三阶偏导存在且有界, f 有紧支集或 K 是有界且有紧支集. 那么

$$H_T\left(\hat{\vartheta}_T-\vartheta(x)-\begin{bmatrix}b_h(x)\\ 0\end{bmatrix}\right)\xrightarrow{d}N(0,A^{-1}(x)V(x)(A^{-1})'),\quad T\to\infty,$$

其中

$$b_h(x)=\sum_{i=1}^{d}h_i^2m_{ii}^{(2)}(x)f(x)\widetilde{L}_2,$$

$$A^{-1}=\begin{bmatrix}1/f(x) & 0\\ -\nabla f(x)/f^2(x) & \mathbf{I}/(\widetilde{L}_2f(x))\end{bmatrix}.\tag{9.10}$$

且

$$\Sigma=\begin{bmatrix}\widetilde{L}_2^d\sigma^2(x)/f(x) & 0\\ 0 & \mathbf{I}_d\,\widetilde{L}_2^{d-1}\,\widetilde{L}_{22}/(\widetilde{L}_2^2f(x))\end{bmatrix},$$

这里 $\widetilde{L}_{22}=\int u^2L^2(u)\mathrm{d}u$.

证明 因为 $\{X_t\}$ 是遍历的, 所以核平滑的弱大数定律(即定理 8.5.8 的(8.23)式, 它可以容易地推广到多元情形)对函数 $\psi(X_t, x)=X_t-x$ 和 $\psi(X_t, x)=(X_t-x)(X_t-x)'$ 成

343

立，下面的证明将经常用到事实

$$\frac{1}{Th}\sum_{t=1}^{T}K\left(\frac{X_t-x}{h}\right)\psi(X_t,x)=\int\frac{1}{h}K\left(\frac{z-x}{h}\right)\psi(z,x)f(z)\mathrm{d}z+o_P(1).$$

由(9.6)式有

$$H_T(\hat{\vartheta}_T(x)-\vartheta(x))=A_T^{-1}(x)\{U_T(x)+H_TB_T(x)+H_TR_T(x)\},\tag{9.11}$$

其中 $U_T(x)$ 由(9.9)式定义，且

$$B_T(x)=\frac{1}{Th_1\cdots h_d}\sum_{t=1}^{T}K\left(\frac{X_t-x}{h}\right)\begin{pmatrix}1\\ \dfrac{X_t-x}{h^2}\end{pmatrix}\frac{1}{2}(X_t-x)'Dm(x)(X_t-x)$$

$$R_T(x)=\frac{1}{Th_1\cdots h_d}\sum_{t=1}^{T}K\left(\frac{X_t-x}{h}\right)\begin{pmatrix}1\\ \dfrac{X_t-x}{h^2}\end{pmatrix}o(\|X_t-x\|^2).$$

下面证明

（ⅰ）$A_T(x)\xrightarrow{P}A(x)(T\to\infty)$ 对某个可逆矩阵 $A(x)$ 成立.

（ⅱ）$H_TB_T(x)=\begin{pmatrix}H_{T1}(b_h(x)+o(1)+o_P(1))\\ H_{T2}\max_j h_j+o_P(1)\end{pmatrix}$，其中

$$H_T=\begin{pmatrix}H_{T1}&0\\0&H_{T2}\end{pmatrix},$$

$$H_{T1}=\sqrt{Th_1\cdots h_d},H_{T2}=\sqrt{Th_1\cdots h_d}\operatorname{diag}(h_1,\cdots,h_d).$$

（ⅲ）$H_TR_T(x)=o_P(1)$.

显然 $R_T(x)$ 的阶比 $B_T(x)$ 低，（ⅲ）是显然的，这里略去其细节.（ⅱ）和（ⅲ）表明

$$H_T\left(\hat{\vartheta}_T(x)-\vartheta(x)-\begin{pmatrix}b_h(x)\\0\end{pmatrix}\right)=A_T^{-1}(x)U_T(x)+o_P(1),\quad T\to\infty.$$

结合（ⅰ）和(9.9)式，由 Slutzky 引理知定理结论成立. 现在证明（ⅰ）. 首先证明序列 $A_T(x)(T\geqslant1)$ 收敛于某个正规矩阵 $A(x)$. 将 $A_T(x)$ 分块为

$$A_T(x)=\begin{pmatrix}\alpha_T(x)&\beta_T(x)\\ \gamma_T(x)&\delta_T(x)\end{pmatrix},$$

344

其中

$$\alpha_T(x)=\frac{1}{Th_1\cdots h_d}\sum_{t=1}^{T}K\left(\frac{X_t-x}{h}\right),$$

$$\beta_T(x)=\frac{1}{Th_1\cdots h_d}\sum_{t=1}^{T}K\left(\frac{X_t-x}{h}\right)(X_t-x)',$$

$$\gamma_T(x)=\frac{1}{Th_1\cdots h_d}\sum_{t=1}^{T}K\left(\frac{X_t-x}{h}\right)\frac{(X_t-x)}{h^2},$$

$$\delta_T(x)=\frac{1}{Th_1\cdots h_d}\sum_{t=1}^{T}K\left(\frac{X_t-x}{h}\right)\frac{(X_t-x)(X_t-x)'}{h^2}.$$

显然，当 $T\to\infty$ 时，$\alpha_T(x)\to f(x)$（依概率）. 我们考虑 $\delta_T(x)$ 的 $(i,\ j)$ 元素. 由常用替换

$u_l=(z_l-x_l)/h_l$，$l=1,\ldots,d$，得到

$$
\begin{aligned}
(\delta_T(x))_{i,j} &= \frac{1}{Th_1\cdots h_d}\sum_{t=1}^{T}K\left(\frac{X_t-x}{h}\right)\frac{X_{ti}-x_i}{h_i}\frac{X_{tj}-x_j}{h_j} \\
&= \int \frac{1}{\prod_{j=1}^{d}h_j}K\left(\frac{z-x}{h}\right)\frac{z_i-x_i}{h_i}\frac{z_j-x_j}{h_j}f(z)\mathrm{d}z+o_P(1) \\
&= \int K(u)u_iu_jf(x+hu)\mathrm{d}u+o_P(1) \\
&= \int K(u)u_iu_j[f(x)+\nabla f(x)hu+(1/2)(hu)'Df(x)(hu)+r_f]\mathrm{d}u+o_P(1).
\end{aligned}
$$

这里以及整个证明，都用到三阶泰勒展开式

$$
f(x+hu)=f(x)+\nabla f(x)hu+(hu)'Df(x)(hu)+r_f, \tag{9.12}
$$

余项为 $r_f=O\left(\sum_{j=1}^{d}h_j^3\right)$，详见(8.21)式及其讨论. 因为

$$
\int u_iu_jK(u)\mathrm{d}u=\begin{cases}\widetilde{L}_2, & i=j,\\ 0, & i\neq j.\end{cases}
$$

进一步

$$
\begin{aligned}
\int K(u)u_iu_j\ \nabla f(x)hu\,\mathrm{d}u &= \sum_{k=1}^{d}h_k\frac{\partial f(x)}{\partial x_k}\int K(u)u_iu_ju_k\mathrm{d}u \\
&= h_i\frac{\partial f(x)}{\partial x_i}\int K(u)u_i^2u_j\mathrm{d}u+h_j\frac{\partial f(x)}{\partial x_j}\int K(u)u_j^2u_i\mathrm{d}u \\
&= \begin{cases}0, & i\neq j,\\ O(\max\{h_i,h_j\}), & i=h.\end{cases}
\end{aligned}
$$

345

由于 $f(x)\int K(u)u_iu_j\mathrm{d}u=f(x)\widetilde{L}_{22}\delta_{ij}$，其中 $\delta_{ij}=1(i=j)$ 为 Kroncker 记号，则得到

$$
(\delta_T(x))_{i,j}=f(x)\widetilde{L}_2+o(1).
$$

因此，

$$
\delta_T(x)=\mathbf{I}_df(x)\widetilde{L}_2+o(1)+o_P(1),\quad T\to\infty.
$$

现考虑 $\gamma_T(x)$的第 k 个元，则有

$$
\begin{aligned}
(\gamma_T(x))_k &= \frac{1}{Th_1\cdots h_d}\sum_{t=1}^{T}K\left(\frac{X_t-x}{h}\right)\frac{X_{tk}-x_k}{h_k^2} \\
&= \frac{1}{h_k}\int \frac{1}{\prod_{j=1}^{d}h_j}K\left(\frac{z-x}{h}\right)\frac{z_{tk}-x_k}{h_k}f(z)\mathrm{d}z+o_P(1) \\
&= \frac{1}{h_k}\int u_kK(u)[f(x)+\nabla f(x)hu+(1/2)(hu)'Df(x)(hu)+r_f]\mathrm{d}u+o_P(1) \\
&= \frac{1}{h_k}\int u_kK(u)[\nabla f(x)hu+(1/2)(hu)'Df(x)(hu)+r_f]\mathrm{d}u+o_P(1),
\end{aligned}
$$

因为 $\int u_k K(u)\mathrm{d}u = 0$，所以含有 $f(x)$的项消失. 进一步

$$O\left(\frac{1}{h_k}\int |u_k K(u)| \mathrm{d}u r_f\right) = O\left(\sum_{j=1}^{d} h_j^2\right) = o(1).$$

接下来讨论泰勒展式中的线性项及二次项. 由

$$\begin{aligned}\frac{1}{h_k}\int u_k K(u)\, \nabla f(x) hu\,\mathrm{d}u &= \sum_{j=1}^{d} \frac{h_j}{h_k} \frac{\partial f(x)}{\partial x_j}\int u_k u_j K(u)\mathrm{d}u \\ &= \frac{\partial f(x)}{\partial x_k}\int u_k^2 K(u)\mathrm{d}u \\ &= \frac{\partial f(x)}{\partial x_k}\widetilde{L}_2\end{aligned}$$

且

$$\frac{1}{2h_k}\int u_k K(u)(hu)' D f(x)(hu)\mathrm{d}u = o(1),$$

则得到

$$\gamma_T(x) = \widetilde{L}_2\, \nabla f(x) + o(1) + o_P(1), \quad T \to \infty.$$

注意到 $\beta_T(x) = h^2 \gamma_T(x)'$，可得

$$\beta_T(x) = o(1) + o_P(1),$$

346

这就证明了(ⅰ)，即

$$A_T \xrightarrow{P} A = \begin{bmatrix} f(x) & 0 \\ \widetilde{L}_2\, \nabla f(x) & \mathbf{I}_d f(x)\, \widetilde{L}_2 \end{bmatrix}.$$

显然，A 是正规的，且它的逆由(9.10)式给出. 最后证明(ⅱ)，考虑 $H_T B_T(x)$的第一项. 由常用替换，有

$$\begin{aligned}(H_T B_T(x))_1 &= \sqrt{T h_1 \cdots h_d}\, \frac{1}{2T}\sum_{t=1}^{T} \frac{1}{\prod_{j=1}^{d} h_j} K\left(\frac{X_t - x}{h}\right)(X_t - x)' D m(x)(X_t - x) \\ &= \sqrt{T h_1 \cdots h_d} \sum_{i,j=1}^{d} m_{ij}^{(2)}(x)\, \frac{1}{T}\sum_{t=1}^{T} \frac{1}{\prod_{k \neq i,j}^{d} h_k} K\left(\frac{X_t - x}{h}\right) \frac{X_{ti} - x_i}{h_i}\, \frac{X_{tj} - x_i}{h_j}, \\ &= \sqrt{T h_1 \cdots h_d} \sum_{i,j=1}^{d} \left\{ m_{ij}^{(2)}(x) h_i h_j \int K(u) u_i u_j f(x + hu)\mathrm{d}u + o_P(1) \right\},\end{aligned}$$

其中 $Dm(x) = (m_{ij}^{(2)}(x))_{i,j}$，最后一个等式成立是因为

$$\begin{aligned}&\int \frac{1}{\prod_{k \neq i,j}^{d} h_k} K\left(\frac{z - x}{h}\right) \frac{z_i - x_i}{h_i}\, \frac{z_j - x_j}{h_j} f(z)\mathrm{d}z \\ &= h_i h_j \int u_i u_j K(u) f(x + hu)\mathrm{d}u.\end{aligned}$$

再一次应用泰勒展式(9.12)，考虑第一项，则有

$$(H_T B_T(x))_1 = \sqrt{Th_1 \cdots h_d} \sum_{i,j=1}^{d} m_{ij}^{(2)}(x) \left\{ h_i h_j f(x) \int u_i u_j K(u) \mathrm{d}u + o_P(1) \right\}$$
$$= \sqrt{Th_1 \cdots h_d} \sum_{i=1}^{d} h_i^2 m_{ii}^{(2)}(x) \{ f(x) \widetilde{L}_2 + o_P(1) \}.$$

现在考虑 $H_T B_T(x)$的第 k 项，$k>1$，有

$$\sqrt{Th_1 \cdots h_d} h_k \frac{1}{T} \sum_{t=1}^{T} \frac{1}{\prod_{l=1}^{d} h_l} K\left\{ \frac{X_t - x}{h} \right\} \frac{X_{tk} - x_k}{h_k^2} \frac{1}{2} \sum_{i,j=1}^{d} (X_{tj} - x_j) m_{ij}^{(2)}(x) (X_{ti} - x_i)$$

$$= \sqrt{Th_1 \cdots h_d} \frac{1}{2} \sum_{i,j=1}^{d} h_i h_j \left\{ \int \frac{1}{\prod_{l=1j}^{d} h_l} K\left(\frac{z-x}{h} \right) \frac{z_k - x_k}{h_k} \frac{z_i - x_i}{h_i} \frac{z_j - x_j}{h_j} f(z) \mathrm{d}z + o_P(1) \right\}$$

$$= \sqrt{Th_1 \cdots h_d} \frac{1}{2} \sum_{i,j=1}^{d} \left\{ h_i h_j \int K(u) u_i u_j u_k \left[f(x) + \nabla f(x) hu + O\left(\sum_l h_l^2 \right) \right] \mathrm{d}u + o_P(1) \right\}$$

$$= \sqrt{Th_1 \cdots h_d} \frac{1}{2} \sum_{i,j=1}^{d} \left\{ h_i h_j \int K(u) u_i u_j u_k \left[\nabla f(x) hu + O\left(\sum_l h_l^2 \right) \right] \mathrm{d}u + o_P(1) \right\}.$$

347

最后，估计含有线性项 $\nabla f(x) hu$ 的项. 注意到：若 $d=1$，则 $i=j=k$，从而得到 $\int u_i u_j u_k K(u) \mathrm{d}u = \int t^3 L(t) \mathrm{d}t = 0$ ，这是因为 $\int_{-K}^{0} t^3 L(t) \mathrm{d}t = -\int_{0}^{K} z^3 L(t) \mathrm{d}t$ 对任意 $K>0$ 成立. 否则，如果 $i \neq j$，则 i 和 j 有一个不等于 k，从而 $\int u_i u_j u_k K(u) \mathrm{d}u = 0$ ，如果 $i=j$，可以找到某个 $1 \leqslant v \leqslant d$ 使得 $v \neq i = j$，同理可以讨论. 所以得到

$$h_i h_j \int u_i u_j u_k K(u) \left(\sum_{l=1}^{d} \frac{\partial f(x)}{\partial x_l} h_l u_l \right) \mathrm{d}u$$
$$= h_i h_j \sum_{l=1}^{d} h_l \frac{\partial f(x)}{\partial x_l} \int u_i u_j u_k K(u) u_l \mathrm{d}u$$
$$= h_i h_j \sum_{l \in \{i,j,k\}} h_l \frac{\partial f(x)}{\partial x_l} \int u_i u_j u_k u_l K(u) \mathrm{d}u$$
$$= O(h_i h_j \max\{h_i, h_j, h_k\})$$
$$= O(\max\{h_i^3, h_j^3\}).$$

归纳上面的讨论，由

$$O\left(\sqrt{Th_1 \cdots h_d} \int |u_i u_j u_k K(u)| \mathrm{d}u \sum_{j=1}^{d} h_j^2 \right) = O\left(\sqrt{Th_1 \cdots h_d \sum_{j=1}^{d} h_j^4} \right),$$

因此得到

$$(H_T B_T(x))_k = \sqrt{Th_1 \cdots h_d} \frac{1}{2} \sum_{i,j=1}^{d} \{ o(1) + o_P(1) \} = \sqrt{Th_1 \cdots h_d} (0 + o_P(1)). \quad ■$$

定理 9.2.4 给出了满足平滑中心极限定理(9.9)的任何时间序列的局部线性估计量的渐近性.

定理 9.2.5 设 $\varepsilon_t = Y_t - E(Y_t | X_t)$ 是一个独立同分布的随机变量序列，满足 $E(\varepsilon_1^2) < \infty$. 则(9.9)式成立，且

$$V = \sigma^2(x) f^2(x) \begin{pmatrix} L_2^d & \mathbf{1}' \widetilde{L}_2 L_2^{d-1} \\ \mathbf{1} \widetilde{L}_2 L_2^{d-1} & W \end{pmatrix},$$

其中 W 是一个 d 维矩阵，其对角元等于 $\widetilde{L}_2 L_2^{d-1}$，非对角元等于 $L_2^{d-1}\left(\int uL^2(u)\mathrm{d}u\right)^2$.

证明 易证定理 8.5.4 对随机向量也成立. 因此只需证明：当 $T \to \infty$ 时

$$V_T = \operatorname{Var}\left(\frac{1}{\sqrt{Th_1 \cdots h_d}} \sum_{t=1}^{T} K\left(\frac{X_t - x}{h}\right) \begin{pmatrix} 1 \\ \dfrac{X_t - x}{h} \end{pmatrix} \varepsilon_t\right)$$

348

收敛于 V. 因为

$$\begin{aligned} V_T &= \frac{1}{Th_1 \cdots h_d} \sum_{t=1}^{T} E\left(K^2\left(\frac{X_t - x}{h}\right) \begin{pmatrix} 1 \\ \dfrac{X_t - x}{h} \end{pmatrix} \left(1, \frac{(X_t - x)'}{h}\right) \varepsilon_t\right) \\ &= \frac{1}{h_1 \cdots h_d} E_{X_1} E\left(K^2\left(\frac{X_1 - x}{h}\right) \begin{pmatrix} 1 \\ \dfrac{X_1 - x}{h} \end{pmatrix} \left(1, \frac{(X_1 - x)'}{h}\right) E(\varepsilon_1^2 | X_1)\right) \\ &= \frac{1}{h_1 \cdots h_d} \int K^2\left(\frac{z - x}{h}\right) \begin{pmatrix} 1 \\ \dfrac{z - x}{h} \end{pmatrix} \left(1, \frac{(z - x)'}{h}\right) \sigma^2(z) f(z) \mathrm{d}z \\ &= \int K^2(u) \begin{pmatrix} 1 \\ u \end{pmatrix} (1, u')(\sigma^2 f)(x + hu) \mathrm{d}u. \end{aligned}$$

注意到 $K^2(u) = \prod_{j=1}^{d} L^2(u_j)$，$u = (u_1, \cdots, u_d)$，使用 Bochner 引理 d 次得到

$$\begin{aligned} \int K^2(u)(\sigma^2 f)(x + hu)\mathrm{d}u &= \int \prod_{j=1}^{d} L^2(u_j)(\sigma^2 f)(x + hu)\mathrm{d}u_1 \cdots \mathrm{d}u_d \\ &= (\sigma^2 f)(x)\left(\int L^2(t)\mathrm{d}t\right)^d + o(1), \end{aligned}$$

注意到 $\int K^2(u)(\sigma^2 f)(x + hu)\mathrm{d}u$ 等于

$$\int \prod_{j=2}^{d} L(u_j)\left[\int L^2(u_1)(\sigma^2 f)(x_1 + h_1 u_1, x_2 + h_2 u_2, \cdots, u_d + h_d u_d)\mathrm{d}u_1\right]\mathrm{d}u_2 \cdots \mathrm{d}u_d,$$

其中，括号中的表达式收敛于 $(\sigma^2 f)(x_1, x_2 + h_2 u_2, \cdots, x_d + h_d u_d)$，$h_1 \to 0$. 如此下去，对 $k = 1, \cdots, d$，

$$\begin{aligned} \int K^2(u) u_k (\sigma^2 f)(x + hu)\mathrm{d}u &= (\sigma^2 f)(x) \int u_k L^2(u_k)\mathrm{d}u_k \prod_{j \neq k} \int L^2(u_j)\mathrm{d}u_j + o(1) \\ &= (\sigma^2 f)(x) \int tL^2(t)\mathrm{d}t \left(\int L^2(t)\mathrm{d}t\right)^{d-1} + o(1), \end{aligned}$$

还有

$$\int K^2(u)u_iu_j(\sigma^2 f)(x+hu)\mathrm{d}u$$
$$=\begin{cases}(\sigma^2 f)(x)\int u_iL^2(u_i)\mathrm{d}u\int u_jL^2(u_j)\left(\int L^2(t)\mathrm{d}t\right)^{d-2}, & i\neq j,\\ (\sigma^2 f)(x)\left(\int L^2(t)\mathrm{d}t\right)^{d-1}\int t^2L(t)\mathrm{d}t, & i=j.\end{cases}$$

349 由$\widetilde{L}_2=\int t^2L(t)\mathrm{d}t, L_2=\int L(t)\mathrm{d}t$，结论可得.

9.3 变点检测和监测

由于来自金融中的时间序列的真实有限维分布是未知的，推断这些分布需要大样本的渐近理论，既有参数方法又有非参数方法. 本书通篇都是假设有了一个无限单边或双边时间序列

$$X_1, X_2, \cdots \quad \text{或} \quad \cdots, X_{-1}, X_0, X_1, \cdots$$

的前 T 个观测 X_1，…，X_T，T 是足够大的，从而算得统计量 $U_T(X_1, \cdots, X_T)$的值. 例如进行一个统计检验，总体的分布由中心极限定理近似估计出，以此来确定一些像临界值，置信区间或风险测度这种人们感兴趣的东西.

通常第 n 个观测值 X_n 是可以通过观测得到的，当然可能会有一点延迟，以下忽略这种延迟. 这意味着在经典的方法中，在 T 时刻要运用推断过程来做决定，之前必须等待了 T 个时间单位. 然而这种决定是可以更早地做出的，而且很多金融任务如交易员的风险监测，代客投资组合管理，流动性管理或通过发起对冲、卖出风险或平仓来处理风险等，其不连续行为在极短的时间内发生.

在这种情况下，监测是有序的，依次分析数据流 X_1，X_2，…. 在时刻 t，可得 t 个已知样本 X_1，X_2，…，X_t，通过计算所谓的控制统计量来进行分析，根据控制统计量的值，人们决定继续监测(不行动)或发出一个警告信号. 分析师、交易员、投资组合或风险经理只有当这样一个连续监测过程发生信号时才参与进来，以便决定如何应对，但是在没有信号发出时，他们可以专注于其他问题. 正因如此，序贯方法变得非常吸引人，因为它可以对大量金融工具、头寸或投资组合进行自动监测.

金融数据的一个重要问题是：真实的时间序列通常是非平稳的，因为它们是金融关系和实体的量化. 充其量，可以假定它们在短时间内是平稳的. 这种情况下，时间序列 X_1，X_2，…可以分解成

$$X_1, \cdots, X_{q_1-1}, \quad X_{q_1}, \cdots, X_{q_2-1}, \quad X_{q_2}, \cdots, X_{q_3-1}, \cdots,$$

其中每一块是弱平稳或严平稳的，但是任何包含两个或者更多块子集的时间序列是非平稳的. 在这种情形，称 q_1，q_2，…，q_L 为**变点**. L 可以是无限的，但是通常只考虑变点数目是有限的情形.

边际与(或)有限维分布的变化，或诱导函数(比如自协方差函数)的变化，可以有很多特定的形式. 基本的均值变化模型假定期望在一个变点改变，使得均值函数 $m(t)=$

$E(X_t)$，$t\geqslant 0$ 是一个分段线性函数，

$$m(t)=\sum_{i=0}^{\infty} m_i \mathbf{1}_{[q_i,q_{i+1})}(t), \tag{9.13}$$

350

其中 m_0，m_1，m_2，…是实数，且 $q_0=0$. 忽略 $t<0$ 时的 X_t 是因为时间序列的这个部分虽然能使数学模型看起来更自然，但是它们在实际中不能观测到. 将没有变化的模型也归入这种模型是自然的，此时只需令 $m_t=m_0$ 对所有的 $t\in\mathbf{N}$，$q_j\in\mathbf{N}_0$ 对所有的 j 成立即可.

值得注意的是：很多其他变点模型可以转化为均值变化模型. 例如，设$\{Z_t\}$是一个零均值时间序列，滞后为 h 的自协方差变点模型为

$$\gamma_Z(h)=E(Z_tZ_{t+h})=\begin{cases}\gamma_0(h), & t<q,\\ \gamma_0(h)+\Delta(h), & t\geqslant q,\end{cases}$$

其中 $\gamma_0(h)$是某个自协方差函数，函数 $\Delta(h)\neq 0$ 使得$(\gamma_0+\Delta)(h)$也是一个自协方差函数，它可以转换为均值变化模型来进行分析，事实上，定义

$$X_t=Z_tZ_{t+h},\quad t\geqslant 1,$$

则有 $E(X_t)=\gamma_Z(h)$对所有的 t 成立.

回到基本的均值变化模型(9.14)上来，注意到模型在 $m_1\neq m_0$ 时可以描述为：在(第一个)变点 $q=q_1$ 后，均值函数是一个分段线性函数. 如果假设均值函数属于一个更一般的函数类则会更接近实际问题些. 变点情形可以规入下列模型，对均值函数 $m(t)=E(Y_t)$，

$$m(t)=m_0\mathbf{1}_{[0,q)}(t)+m^*(t)\mathbf{1}_{[q,\infty)}(t), \tag{9.14}$$

其中 m^* 是某个函数，满足 $m^*(q+)\neq m_0$.

9.3.1 离线检测

给定样本 X_1，…，X_T，目标是检验是否存在一个变点，在这个变点处观测值的(边际)分布发生了变化. 在研究均值变化问题前，先简要地讨论一下一般情形. 设

$$X_t\sim f_0(x),\quad t=1,\cdots,k-1,$$
$$X_t\sim f_1(x),\quad t=k,\cdots,T,$$

其中，f_0 和 f_1 是两个不同的密度. 下面假设 f_0 和 f_1 是 Lebesgue 密度函数，但得出的结论能容易地推广到被一个测度 μ 控制的一般密度上去. 与前面一样，设 k 是固定的非随机未知变点. 该问题的**似然比统计量**为：在集合 $A=\left\{\prod_{t=1}^{T}f_0(X_t)\neq 0\right\}$ 上

$$\Lambda_T=\Lambda_T(k)=\frac{\prod_{t=1}^{k-1}f_0(X_t)\prod_{t=k}^{T}f_1(X_t)}{\prod_{t=1}^{T}f_0(X_t)}=\prod_{t=k}^{T}\frac{f_1(X_t)}{f_0(X_t)},$$

在 A^C 上 $\Lambda_T=0$. 注意到 $P_0(A)=1$,，其中 P_0 表示对所有的 t，$X_t\overset{\text{i.i.d.}}{\sim}f_0$ 而求得的概率.

351

如果 Λ_T 取得较大的数，则似然比(LR)检验拒绝原假设

$$H_0: k=\infty(\text{没有变化})$$

而接受 $k>0$ 时刻变点的备择假设，即相应的统计检验为

$$\phi_{LR}=\mathbf{1}(\Lambda_T>c),$$

对某个临界值 c. 现在的问题是：ϕ_{LR} 是否在 $P_0(\phi=1)$ 有相同显著水平的所有统计检验 ϕ 中是最优的？即势 $P_k(\phi=1)$ 是最大的？其中 P_k 表示假设在时刻 $k\in\{1,\cdots,T\}$ 变点计算得到的概率.

下列关于测度变化的引理是重要的.

引理 9.3.1 设 A 是一个可测集. 那么

$$E_0(\Lambda_T\mathbf{1}_A)=E_k(\mathbf{1}_A)=P_k(A)$$

且

$$E_0(\Lambda_TY)=E_k(Y),$$

对任意关于 $\sigma(X_1,\cdots,X_T)$ 可测的随机变量 Y 成立，其中 E_0 表示在 $X_1,\cdots,X_T\sim f_0$ 的假定下求期望.

证明 因为

$$\begin{aligned}E_0(\Lambda_T\mathbf{1}_A)&=\int_A\prod_{t=k}^{T}\frac{f_1(x_t)}{f_0(x_t)}\prod_{t=1}^{T}f_0(x_t)\mathrm{d}x_1\cdots\mathrm{d}x_T\\&=\int_A\prod_{t=1}^{k-1}f_0(x_t)\prod_{t=k}^{T}f_1(x_t)\mathrm{d}x_1\cdots\mathrm{d}x_t\\&=E_k(\mathbf{1}_A)\\&=P_k(A).\end{aligned}$$

类似地，

$$E_0(\Lambda_TY)=\int\prod_{t=1}^{k-1}f_0(x_t)\prod_{t=k}^{T}f_1(x_t)Y\mathrm{d}x_1\cdots\mathrm{d}x_T=E_k(Y).$$ ■

下面是 LR 检验中最优性的一个简洁证明.

定理 9.3.2 设 $\delta:(\Omega,\mathcal{F},P)\to([0,1],\mathcal{B}_{[0,1]})$ 是任意一个随机统计检验. 若 ϕ_{LR} 和 δ 有相同的显著水平，即

352

$$P_0(\Lambda_T>c)=P_0(\delta=1)=\alpha,$$

则 ϕ_{LR} 优于 δ，即

$$P_k(\Lambda_T>c)\geqslant P_k(\delta=1).$$

证明 因为 δ 在 $[0,1]$ 上取值，所以对任意 $x,y\in\mathbf{R}$，有

$$(x-y)[\mathbf{1}(x>y)-\delta]\geqslant 0.$$

在上面的不等式中，取 $x=\Lambda_T$，$y=c$，得到

$$(\Lambda_T-c)\mathbf{1}(\Lambda_T>c)\geqslant(\Lambda_T-c)\delta. \tag{9.15}$$

这表明

$$E_0((\Lambda_T-c)\mathbf{1}(\Lambda_T>c))\geqslant E_0(\Lambda_T\delta)-E_0(c\delta),$$

即

$$E_0(\Lambda_T\mathbf{1}(\Lambda_T>c))-cP_0(\Lambda_T>c)\geqslant E_0(\Lambda_T\delta)-cP_0(\delta=1).$$

由前一个引理知，

$$E_0(\Lambda_T\delta) = E_k(\delta) = P_k(\delta = 1)$$

且

$$E_0(\Lambda_T \mathbf{1}(\Lambda_T > c)) = E_k(\phi_{LR}) = P_k(\phi_{LR} = 1).$$

因此，(9.15)式等价于

$$P_k(\phi_{LR} = 1) - cP_0(\Lambda_T > c) \geqslant P_k(\delta = 1) - cP_0(\delta = 1),$$

由假设 $P_0(\phi_{LR}=1)=P_0(\delta=1)$，所以得到

$$P_k(\phi_{LR} = 1) \geqslant P_k(\delta = 1).$$

■

LR 检验需要知道变化前和变化后的分布 f_0 和 f_1，否则无法计算检验统计量，但实际中这种分布常常是不知道的. 一个常用的办法是对 LR 检验统计量作一些特定的假定，如假设观察值服从某个理想分布(比如正态分布)，然后找到一个适当的极限定理使得结论对非正态观测值是渐近有效的. 这样，比如要求临界值或者得到关于势的结论，就可以通过对统计量进行近似计算得到. 最后，还要考虑这种方法是否可以推广到相依时间序列，可能要作适当的修正才行，这在金融数据分析中是特别值得关注的.

接下来，将集中研究均值变化问题，并先假设 X_t 是独立正态变量，且记均值 $\mu_t = E(X_t)$，方差 $\sigma^2 \in (0, \infty)$. 则均值变化变点检验问题的原假设为

353

$$H_0: \mu_1 = \cdots = \mu_T,$$

备择假设为

$$H_1: \mu_1 = \cdots = \mu_{k-1} \neq \mu_k = \cdots = \mu_T, \quad \text{对某个 } 1 < k < T.$$

如果存在一个变化，它发生在某个固定的时刻 k，考虑检验问题

$$H_0: \text{没有变化} \quad \longleftrightarrow \quad H_1^{(k)}: \text{在时刻 } k \text{ 变化}.$$

对应的似然比统计量为

$$\Lambda_k = \frac{L_k(X_1, \cdots, X_T)}{L_1(X_1, \cdots, X_T)},$$

其中

$$L_k(X_1, \cdots, X_T) \sup_{\mu_1, \mu_2 \in \mathbf{R}} \prod_{t=1}^{k-1} \varphi(X_t - \mu_1) \prod_{t=k}^{T} \varphi(X_t - \mu_2)$$

是备择假设的似然函数，它是对变化前均值 μ_1 和变化后均值 μ_2 和取上确界后得到的 X_1，…，X_T 的联合密度函数，且

$$L_1(X_1, \cdots, X_T) = \sup_{\mu \in \mathbf{R}} \prod_{t=1}^{T} \varphi(X_t - \mu)$$

是原假设的似然函数. 遍及全书，$\varphi(x) = \varphi_{(\mu,\sigma)}(x)$ 均表示 $N(\mu, \sigma^2)$ 分布的密度函数.

因为求乘积

$$\prod_{t=a}^{b} \varphi(X_t - \mu) = \left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^{(b-a+1)} \exp\left\{-\frac{1}{2\sigma^2} \sum_{t=a}^{b} (X_t - \mu)^2\right\},$$

$1 \leqslant a \leqslant b \leqslant T$，$\mu \in \mathbf{R}$，关于 μ 的最大值点等价于对 $\sum_{t=a}^{b} (X_t - \mu)^2$ 求最小值点，所以由最小

二乘法可得它的唯一最大值点为$\frac{1}{b-a+1}\sum_{t=a}^{b}X_t$. 记

$$\overline{X}_k=\frac{1}{k}\sum_{t=1}^{k}X_t,\quad \overline{X}_{(k+1):T}=\frac{1}{T-k}\sum_{t=k+1}^{T}X_t,$$

则得到

$$\Lambda_k(X_1,\cdots,X_T)=\frac{\prod_{t=1}^{k}\varphi(X_t-\overline{X}_k)\prod_{t=k+1}^{T}\varphi(X_t-\overline{X}_{(k+1):T})}{\prod_{t=1}^{T}\varphi(X_t-\overline{X}_T)}.$$

接下来，考虑对数似然比统计量

$$l_k(X_1,\cdots,X_T)=\log\Lambda_k(X_1,\cdots,X_T).$$

引理 9.3.3 对数似然比统计量有下列表达式

354

$$l_k(X_1,\cdots,X_T)=\frac{1}{2\sigma^2}\frac{T}{(T-k)k}\Big(S_k-\frac{k}{T}S_T\Big)^2$$

且

$$l_k(X_1,\cdots,X_T)=\frac{1}{2\sigma^2}\frac{T}{(T-k)k}\Big(\sum_{t=1}^{k}(X_t-\overline{X}_T)^2\Big)^2,$$

其中

$$S_k=\sum_{t=1}^{k}X_t,\quad S_{(k+1):T}=\sum_{t=k+1}^{T}X_t.$$

证明 证明过程涉及一些计算技巧. 首先，取对数后，含有$\sqrt{2\pi\sigma}$的因子可以不考虑了，所以 $l_k(X_1, \cdots, X_T)$等于

$$\frac{1}{2\sigma^2}\Big\{-\sum_{t=1}^{k}(X_t-\overline{X}_k)^2-\sum_{t=k+1}^{T}(X_t-\overline{X}_{(k+1):T})^2+\sum_{t=1}^{T}(X_t-\overline{X}_T)^2\Big\}.\tag{9.16}$$

由简单等式

$$2S_{(k+1):T}\overline{X}_{(k+1):T}=\frac{2}{T-k}S^2_{(k+1):T}\quad 和\quad k\overline{X}_k^2=\frac{1}{k}S_k^2$$

合并后，(9.16)式等于

$$\frac{1}{2\sigma^2}\Big\{\Big(\frac{2}{k}-\frac{1}{k}\Big)S_k^2+\Big(\frac{2}{T-k}-\frac{1}{T-k}\Big)S^2_{(k+1):T}-\frac{1}{T}S_T^2\Big\}.$$

再由 $S^2_{(k+1):T}=(S_T-S_k)^2=S_T^2-2S_kS_T+S_k^2$，合并后得到

$$\begin{aligned}l_k(X_1,\cdots,X_T)&=\frac{1}{2\sigma^2}\Big\{\Big(\frac{1}{k}+\frac{1}{T-k}\Big)S_k^2-\Big(\frac{1}{T}-\frac{1}{T-k}\Big)S_T^2-\frac{2}{T-k}S_kS_T\Big\}\\&=\frac{1}{2\sigma^2}\Big\{\frac{T}{(T-k)k}S_k^2-\frac{k}{(T-k)k}S_T^2-\frac{2}{T-k}S_kS_t\Big\}\\&=\frac{1}{2\sigma^2}\frac{T}{(T-k)k}\Big\{S_k^2-\frac{k}{T}S_T^2-2\,\frac{k}{T}S_kS_T\Big\}.\end{aligned}$$

因此

$$l_k(X_1,\cdots,X_T)=\frac{1}{2\sigma^2}\frac{T}{(T-k)k}\Big(S_k-\frac{k}{T}S_T\Big)^2.$$

由此得到了**最大选择的加权 CUSUM 统计量**

$$C_T^w=\max_{1\leqslant k\leqslant T}\sqrt{\frac{T}{k(T-k)}}\sum_{t=1}^{k}(X_t-\overline{X}_T),$$

未加权形式为

$$C_T=\max_{1\leqslant k\leqslant T}\sum_{t=1}^{k}(X_t-\overline{X}_T).$$

355

当观测值是非正态时上面两个统计量也可以使用. 事实上, 对任意满足泛函中心极限定理(FCLT)的时间序列$\{X_t\}$, 极限定律得到的极限是一个布朗运动的泛函. 作为准备, 考虑下列引理. ■

引理 9.3.4 假设$\{X_t\}$满足 FCLT, 即当 $T\to\infty$时,

$$S_T(u)=\frac{1}{\sqrt{T}}\sum_{t=1}^{\lfloor Tu\rfloor}X_t,\quad u\in[0,1]$$

弱收敛于 $\eta B(u)$, 其中, 对 $\eta\in(0,\infty)$是某个常数, B 是标准布朗运动. 那么

$$B_T(u)=\frac{1}{\sqrt{T}}\sum_{t=1}^{\lfloor Tu\rfloor}(X_t-\overline{X}_T)\Rightarrow\eta B^0(u),\quad T\to\infty,$$

其中

$$B^0(t)=B(t)-tB(1),\quad t\in[0,1],$$

是一个布朗桥.

证明 注意到

$$\frac{1}{\sqrt{T}}\sum_{t=1}^{\lfloor Tu\rfloor}(X_t-\overline{X}_T)=\frac{1}{\sqrt{T}}\sum_{t=1}^{\lfloor Tu\rfloor}X_t-\frac{\lfloor Tu\rfloor}{T}\frac{1}{\sqrt{T}}\sum_{t=1}^{T}X_t$$
$$=S_T(u)-\frac{\lfloor Tu\rfloor}{T}S_T(1).$$

由连续映射定理, 最后一个表达式弱收敛于 $\eta B(u)-u\eta B(u)=\eta B^0(u)$. ■

在引入(加权)CUSUM 统计量的一般极限定理时, 引理 9.3.4 仅是一小步.

定理 9.3.5 设$\{X_t\}$满足 FCLT, 即当 $T\to\infty$时,

$$S_T(u)=\frac{1}{\sqrt{T}}\sum_{t=1}^{\lfloor Tu\rfloor}X_t,\quad u\in[0,1]$$

弱收敛于 $\eta B(u)$, 其中 $\eta\in(0,\infty)$是某个常数, B 是标准布朗运动. 那么, 当 $T\to\infty$时

$$C_T^w\Rightarrow\sup_{u\in[0,1]}\frac{\eta B^0(u)}{\sqrt{u(1-u)}},$$

且

$$C_T\Rightarrow\sup_{u\in[0,1]}\eta B^0(u).$$

356

证明 先证 C_T^w 的结论. 注意到: 当 $u\in[0,1/T)$时, $C_T^w(u)=0$. 因此,

$$C_T^w = \max_{1\leqslant k\leqslant T}\sqrt{\frac{T}{k(T-k)}}\sum_{t=1}^{k}(X_t-\overline{X}_T)$$
$$= \sup_{u\in[1/T,1]}\sqrt{\frac{T}{\lfloor Tu\rfloor(T-\lfloor Tu\rfloor)}}\sum_{t=1}^{\lfloor Tu\rfloor}(X_t-\overline{X}_T)$$
$$= \sup_{u\in[1/T,1]}\frac{B_T^0(u)}{g_T(u)},$$

其中

$$g_T(u) = \sqrt{\frac{\lfloor Tu\rfloor}{T}\frac{T-\lfloor Tu\rfloor}{T}}$$

且

$$B_T^0(u) = \frac{1}{\sqrt{T}}\sum_{t=1}^{\lfloor Tu\rfloor}(X_t-\overline{X}_T)$$

是引理 9.3.4 中的过程．当 $T\to\infty$时，有 $B_T^0\Rightarrow\eta B^0$，且

$$\lim_{T\to\infty}\sup_{u\in[0,1]}|g_T(u)-g(u)| = 0,$$

其中

$$g(u) = \sqrt{u(1-u)},\quad u\in[0,1]$$

是一个确定性的函数．因此，当 $T\to\infty$时，$(B_T^0,\ g_T)\Rightarrow(\eta B^0,\ g)$在乘积空间 $D([0,\ 1];\ \mathbf{R})\otimes D([0,\ 1];\ \mathbf{R})$．由连续映射定理知，当 $T\to\infty$时

$$\frac{B_T^0}{g_T}\Rightarrow\frac{\eta B^0}{g},$$

在 $D([0,\ 1];\ \mathbf{R})$中，因为 $B^0(0)=B^0(1)=0$(a. s.)，从而 $\eta B^0/g$ 在 $u\in\{0,\ 1\}$上有定义．再一次应用连续映射定理得到

$$\sup_{u\in[0,1]}\frac{B_T^0(u)}{g_T(u)}]\Rightarrow\sup_{s\in[0,1]}\frac{\eta B^0(u)}{g(u)},\quad T\to\infty,$$

证毕． ■

在实践中，参数 η(即长期方差)是未知的．然后使用样本 X_1，…，X_T，可用 Newey-West 估计量$\hat{\eta}_T$ 估计 η，见(8.28)式．如果定理 8.8.6 的条件满足，可以将统计量 C_T 和C_T^w 分别除以$\hat{\eta}_T$，然后应用上面的极限过程的 $\eta=1$ 的情形．换句话说，当 $T\to\infty$时

357

$$\frac{C_T^w}{\hat{\eta}_T}\Rightarrow\sup_{u\in[0,1]}\frac{B^0(u)}{\sqrt{u(1-u)}},$$

且

$$\frac{C_T}{\hat{\eta}_T}\Rightarrow\sup_{u\in[0,1]}B^0(u),$$

下面研究在备择假设为存在变点 $q=[T\vartheta]$时的上述检验问题，其中均值从 μ 变为 $\mu+\Delta$，$\Delta\neq0$ 是某个常数．现观测值 X_1，X_2，…满足的模型为

$$X_t = \mu+\Delta\mathbf{1}_{\{q,q+1,\cdots\}}(t)+\varepsilon_t,\quad t=1,2,\cdots,\tag{9.17}$$

其中$\{\varepsilon_t\}$是一个零均值的时间序列，满足 FCLT．前面已经看到，引理 9.3.4 定义的 $B_T(u)$

规范了统计量 C_T 和 C_T^w. 下列命题揭示 $B_T(u)$的性质.

命题 9.3.6 设$\{\varepsilon_t\}$满足 FCLT，即当 $T\to\infty$时，

$$S_T^\varepsilon(u)=\frac{1}{\sqrt{T}}\sum_{t=1}^{\lfloor Tu\rfloor}\varepsilon_t,\quad u\in[0,1]$$

弱收敛于 $\eta B(u)$，其中 $\eta\in(0,\infty)$是某个常数，B 是标准布朗运动. 那么，在模型(9.17)中，过程

$$B_T(u)=\frac{1}{\sqrt{T}}\sum_{t=1}^{\lfloor Tu\rfloor}(X_t-\overline{X}_T),\quad u\in[0,1],$$

满足

$$B_T(u)-\delta_T(u)\Rightarrow B^0(u),\quad T\to\infty,$$

其中 B^0 是一个布朗桥，当 $T\to\infty$时，如果 $u<\vartheta$，则漂移项

$$\delta_T(u)=\Delta\frac{\lfloor Tu\rfloor-\lfloor T\vartheta\rfloor+1}{\sqrt{T}}\mathbf{1}(\lfloor Tu\rfloor>\lfloor T\vartheta\rfloor)$$

收敛于 0；如果 $u\geqslant\vartheta$，则漂移项发散到 $\mathrm{sign}(\Delta)\infty$.

证明 注意到

$$\overline{X}_T=\bar{\varepsilon}_T+\Delta\frac{T-q+1}{T}$$

且

$$X_t-\overline{X}_T=\begin{cases}\varepsilon_t-\bar{\varepsilon}_T-\Delta\dfrac{T-q+1}{T}, & t<q,\\[2ex] \varepsilon_t-\bar{\varepsilon}_T+\Delta(t-q+1)-\Delta\dfrac{T-q+1}{T}, & t\geqslant q.\end{cases}$$

定义

$$B_T^\varepsilon(u)=\frac{1}{\sqrt{T}}\sum_{t=1}^{\lfloor Tu\rfloor}(\varepsilon_t-\bar{\varepsilon}_T),\quad u\in[0,1].$$

则有 358

$$\begin{aligned}B_T(u)&=B_T^\varepsilon(u)+\frac{1}{\sqrt{T}}\sum_{t=1}^{\lfloor Tu\rfloor}\left[\Delta(t-q+1)\mathbf{1}_{\{q,q+1,\cdots\}}(t)-\frac{T-q+1}{T}\Delta\right]\\&=B_T^\varepsilon(u)+\Delta\frac{\lfloor Tu\rfloor-q+1}{\sqrt{T}}\mathbf{1}(\lfloor Tu\rfloor\geqslant\lfloor T\vartheta\rfloor)-\frac{T-q+1}{\sqrt{T}}\Delta\\&=B_T^\varepsilon(u)+\delta_T(u)+o(1).\end{aligned}$$

因此，当 $T\to\infty$时

$$B_T(u)-\delta_T(u)=B_T^\varepsilon(u)+o(1)\Rightarrow\eta B(u),$$

这是因为，当 $T\to\infty$时，$S_T^\varepsilon\Rightarrow\eta B$，能推出 $B_T^\varepsilon\Rightarrow\eta B^0$，$T\to\infty$(由引理 9.3.4). ■

命题 9.3.6 表明过程 B_T 在变化前，即$\left\{B_T(u):u<\dfrac{\lfloor Tu\rfloor}{T}\right\}$收敛于布朗桥对应的部分$\{\eta B_T^0: u\in[0,\vartheta]\}$，在变点后过程 $B_T(u)$发散，从而，最大选择的 CUSUM 统计量也发散.

9.3.2 在线检测

上节介绍的离线检测是一次性地分析 T 个样本检测变点. 在金融实践中出现了许多连续检测，也称在线检测或在线监测的方法. 在连续检测中，设观测值是依次取得的，由此形成一个数据流，且在 t 时刻，观察值 X_t 是已知的，即没有延迟或延迟是可忽略不计的. 序贯法在时刻 t 分析已有的 t 个样本 X_1，…，X_t 从而形成一个决策. 实践中常见的方法是按下列方式设计一种决策(检测方法). 首先计算控制统计量 $U_t=U_t(X_1, \cdots, X_t)$，它的取值是一个实数，如果控制统计量的值落在一个临界集 A 中，则发出一个数据支持的建立在经验证据基础上的变点信号；否则监测继续进行到下一时刻. 方法设计时有两种选择，一种是可以永远持续下去的监测方法；另一种是当监测到达某个时间点 T(定义为最大样本数)时则停止的监测方法. 在金融应用中，后一种方法通常是更加自然的，因为基本的假定和模型通常都是定期(比如一个季度或者一年)更新的. 用数学语言描述，$U_t\in A$ 的首达时刻 $t\leqslant T$ 由停时给出，

$$R_T = \inf(k \leqslant t \leqslant T: U_t \in A).$$

下面，设 $T<\infty$.

有非常多形式的控制统计量，下面考虑加权平均形式

359

$$\hat{m}_t = \sum_{i=1}^{t} K\left(\frac{t-i}{h}\right) Y_i, \quad t = 1,2,\cdots,$$

其中 $K(x)$是一个平滑核，$h>0$ 是带宽. 这个统计量与 Nadaraya-Watson 估计量非常相似. 然而，为了检测均值的变化，不需对均值进行一致估计，因此，不需规范权重. 进一步，假设观测值 Y_1，Y_2，…是依次取到的，并且是在等距时间内观测的，在缩放处理后可以假定观测时间为自然数，又假设时间 T 趋于无穷时，带宽趋于无穷. 确切地，设 $h=h_T$ 满足

$$\lim_{T\to\infty} \frac{T}{h_T} = \zeta,$$

其中 $\zeta\in[1, \infty)$. 由于$\hat{m}_T$ 以$\sqrt{T}$速率收敛，考虑停时 R_T，其中 $U_t=\frac{1}{\sqrt{T}}\hat{m}_T$，即

$$R_T = \inf\left\{k \leqslant t \leqslant T: \frac{1}{\sqrt{T}} \frac{1}{\sqrt{T}} \hat{m}_t > c\right\}.$$

定义对应的**顺序加权核部分和过程**为

$$M_T(s) = \sum_{t=1}^{\lfloor Ts \rfloor} K\left(\frac{\lfloor Ts \rfloor - t}{h}\right) Y_t, \quad s \in [0,1],$$

则

$$R_T = T \inf\left\{s \in [0,1]: \frac{1}{\sqrt{T}} M_T(s) > c\right\}.$$

接下来，对均值变化模型(9.14)稍做修正. 设时间 $T\in \mathbf{N}$ 是给定的，用于检测的前 T 个观测值均依赖于 T，即 $Y_t=Y_{Tt}$. 也即假设顺序观测值 Y_{T1}，…，Y_{TT} 满足模型

$$Y_{Tt} = \frac{1}{\sqrt{T}} m_0\left(\frac{t-q}{T}\right) \mathbf{1}(t \geqslant q) + \varepsilon_t, \quad t = 1,\cdots,T, \tag{9.18}$$

其中，$m_0:\mathbf{R}\to[0,\infty)$是某个函数，$q$为变点，$\{\varepsilon_t\}$是一个零均值的严平稳过程.

若函数m_0满足下列条件：存在某个$t^*>0$，

$$m_0(s)=0,\quad s\leqslant 0,\quad 和\ m_0(s)>0, s\in(0,t^*),\tag{9.19}$$

则q是一个变点. 在变点前观测值为零均值的，而在变点后均值函数由严格正函数$m_0(t)$，$t>0$导出. 模型(9.18)是一个局部替代模型：当T不断变大时，变点检测问题变得越来越困难，因为此时均值趋于0. 特别地，如果m_0依上确界范数有界，则$E(Y_{Tt})=O(1/\sqrt{T})$，$t\geqslant q$.

360

下面的结论给出了原假设(没有变点)下的一些分布.

定理 9.3.7 设$\{\varepsilon_t\}$满足$E|\varepsilon_1|^{r+\delta}<\infty$，$r\geqslant 4$，$\delta>0$，且$\{\varepsilon_t\}$是$\alpha$-混合的，混合系数$\alpha(k)$，$k\in\mathbf{N}$满足

$$\alpha(k)=O(k^{-\beta}),对某个\ \beta>\frac{r(r+\delta)}{2\delta}.$$

其中，K为有界的 Lipschitz 连续核. 那么，在没有变点原假设$H_0:m_0=0$下，下列结论成立：

(i) 对所有的$0\leqslant s$，$t\leqslant 1$，极限

$$C(r,s)=\lim_{T\to\infty}\frac{1}{T}\sum_{i=1}^{\lfloor Tr\rfloor}\sum_{j=1}^{\lfloor Ts\rfloor}K\Big(\frac{\lfloor Ts\rfloor-i}{h}\Big)K\Big(\frac{\lfloor Tr\rfloor-j}{h}\Big)r_0(|i-j|)\tag{9.20}$$

存在，其中$r_0(k)=E(\varepsilon_1\varepsilon_{1+k})$，$k\in\mathbf{N}$是$\{\varepsilon_t\}$的自协方差函数.

(ii) 经验过程$\{M_T(s):s\in[0,1]\}$弱收敛

$$M_T\Rightarrow M,\quad T\to\infty$$

其中M是一个零均值的高斯过程，其协方差函数为(9.20).

(iii) 停时R_T满足中心极限定理

$$\frac{R_T}{T}\xrightarrow{d}\inf\{s\in[0,1]:M(s)>c\},\quad T\to\infty. \qquad ■$$

下面研究均值变化条件(9.19)成立时的一些性质. 设变点q为T的函数，即$q=q_T$.

命题 9.3.8 设均值变化条件(9.19)和定理9.3.7的假定均成立. 如果乘积Km_0是可积的，则当$T\to\infty$时，

$$\frac{1}{\sqrt{T}}M_T(s)\Rightarrow M^1(s)=M(s)+\int_{\lim q/T}^{s}K(\zeta(s-r))m_0(r-\lim q/T)\mathrm{d}r,$$

于是，当$T\to\infty$时

$$\frac{R_T}{T}\xrightarrow{d}\inf\{s\in[0,1]:M^1(s)>c\}.$$

证明 结论可由 Slutzky 引理以及下列事实

$$\frac{1}{\sqrt{T}}M_T(s)=\frac{1}{\sqrt{T}}M_T^0(s)+m_T(s)$$

361

得到，其中

$$M_T^0(s)=\frac{1}{\sqrt{T}}\sum_{t=1}^{\lfloor Ts\rfloor}K\Big(\frac{\lfloor Ts\rfloor-t}{h}\Big)\varepsilon_t,$$

$$m_T(s)=\frac{1}{T}\sum_{t=1}^{\lfloor Ts\rfloor}K\Big(\frac{\lfloor Ts\rfloor-t}{h}\Big)m_0\Big(\frac{t-q}{T}\Big)\mathbf{1}(t\geqslant q).$$

根据定理 9.3.7，有 $M_T^0\Rightarrow M$，$T\to\infty$. 进一步，由 K 和 m_0 的 Lipschitz 连续性，当 $T\to\infty$时

$$\begin{aligned}m_T(s)&=\frac{1}{T}\sum_{t=1}^{\lfloor Ts\rfloor}K\Big(\frac{T}{h}\frac{\lfloor Ts\rfloor-t}{T}\Big)m_0\Big(\frac{t-q}{T}\Big)\mathbf{1}(t\geqslant q)\\&=\frac{1}{T}\sum_{t=q}^{\lfloor Ts\rfloor}K\Big(\zeta\Big(s-\frac{t}{T}\Big)\Big)m_0\Big(\frac{t}{T}-\lim\frac{q}{T}\Big)+o(1)\\&=\int_{\lim q/T}^{s}K(\zeta(s-r))m_0(r-\lim q/T)\mathrm{d}r+o(1),\end{aligned}$$

由连续映射定理可得 R_T/T 的结论，见定理 B.2.1 和例 B.2.2(iv). ■

评注 9.3.9 可以看出：极限过程 M^1 通过下列依赖于核 K 的确定性漂移函数

$$\mu(s;K)=\int_{\lim q/T}^{s}K(\zeta(s-r))m_0(r-\lim q/T)\mathrm{d}r,$$

依赖于函数 m_0，而 m_0 确定了变点后过程的均值函数. 一个很自然的问题是，如何在某种合理意义下，用延迟过程 R_T-q 来选取最优核. 下列修正方法，给出了寻找最优核的方法. 考虑模型

$$Y_{Tt}=\frac{1}{\sqrt{T}}m_0\Big(\frac{t-q}{T}\Big)\mathbf{1}(t\geqslant q)+\varepsilon_t,\quad t=1,\cdots,T,$$

其中，$\lim q/T=0$. 如果用控制统计量 $U_t=\hat{m}_t/h$，并考虑停时

$$R_\infty=\inf\{t\geqslant 0:\hat{m}_t/h>c\},$$

则当 $h\to\infty$时，标准化的延迟 $\rho_h=\dfrac{R_T-q}{h}$几乎必然收敛于

$$\rho_0=\inf\{s>0:\widetilde{\mu}(s;K)>c\},$$

其中，$\widetilde{\mu}(s;K)$是某个函数，其定义类似于 $\mu(s;K)$. 现在，一大类函数 m_0 的最优核可以在相当弱的条件下得到，见参考文献. ■

362

9.4 单位根和随机游动

在本节的开始，设 u_1，u_2，…是独立同分布的随机变量，均值为 0，方差为 $\sigma^2=E(u_1^2)\in(0,\infty)$. 其后，将大幅度放宽这一条件. 回忆一下 Bachelier 的股票价格模型，考虑纯随机游动

$$Y_t=\sum_{i=0}^{t}u_i,\quad t=0,1,\cdots,$$

其初始值为 $Y_0=u_0$. 上述 Y_t 的方程可以写成递推形式

$$Y_t=Y_{t-1}+u_t,\quad t=1,2,\cdots,$$

初始值 $Y_0=u_0$. 显然，这一递推形式是下列这类递推形式的一种

$$Y_t = \rho Y_{t-1} + u_t, \quad t = 1,2,\cdots, \tag{9.21}$$

参数 $\rho \in (-1, 1]$. 取 $\rho=1$ 就是上面的纯随机游动，而 $|\rho|<1$ 时，则得到 AR(1)过程，其初始值 $Y_0=u_0$. 在第 3 章中已经讨论了方程(9.21)是否存在平稳解这一问题. 如果 u_{-n}，($n=0$, 1, …)是与 u_1 同分布的 i. i. d.，且 $Y_0=\sum_{i=0}^{\infty}\rho^i u_{-i}$，那么

$$Y_t = \sum_{i=0}^{\infty}\rho^i u_{t-i}, \quad t \in \mathbf{N}_0$$

是 AR(1)方程一个平稳(或因果)解.

一个过程是平稳过程还是一个随机游动可以通过特征多项式

$$\varphi(z) = 1 - \rho z, \quad z \in \mathbf{C}$$

来判断，特征多项式恰有一个根 $z=1/\rho$. 这个根位于单位圆 $\{z\in\mathbf{C}: |z|\leqslant 1\}$ 外当且仅当 $|\rho|<1$. 如果 $\rho=1$，则特征多项式的根在单位圆上. 在单位圆上的根称为**单位根**. 模型(9.21)是一个随机游动的充分必要条件是特征多项式有单位根.

在下面将会看到，当基础时间序列是一个带独立扰动项的平稳 AR(1)或者是一个带相依扰动项的非平稳随机游动时，它们的统计性质完全不同. 对于许多表示对数价格或利率的时间序列，经济学理论和统计理论表明随机游动的假设是有道理的. 统计学的任务是检验这个假设是否可以被已有数据证实. 这就导出了下面将要讨论的关于 AR(1)模型的两个检验问题.

平稳性检验：平稳性检验的原假设为

$$H_0: |\rho| < 1,$$

存在一个单位根的备择假设(随机游动假设)为

$$H_1: \rho = 1.$$

363

单位根检验：单位根检验的原假设与平稳性检验相反，为

$$H_0: \rho = 1,$$

与平稳性相对的备择假设为

$$H_1: |\rho| < 1.$$

在本章的最后，将对这两个检验作推广. 单位根检验在交易中也有有趣的应用，下面是一个例子.

▶**例 9.4.1**(配对交易) 设 $\{X_t\}$ 和 $\{Y_t\}$ 是两个随机游动，例如两个相似资产或商品的价格. 有些资产和商品分别是可交换的，比如可以用黑胡椒交替白胡椒，或者反之. 交换的前提是它们的价格非常相似，即交换物品价格的差值在 0 附近平稳波动. 交换有下列交易策略：如果白胡椒的价格远低于黑胡椒，可以买入多头白胡椒并且卖出空头黑胡椒. 当白胡椒价格变得高于黑胡椒时，恢复原有头寸. 如果差价 $D_t=Y_t-X_t$($t=1$, 2, …)是一个零均值的平稳过程而不是一个随机游动，这个策略就有效. 如果差价是一个随机游动，价差超过一个给定阈值的概率要高于一个平稳过程价差超过同一阈值的概率. 为了验证是否可以进行配对交易，可以假设 D_t 服从模型(9.21)，然后应用单位根检验. 如果检验拒绝是随机游动的原假设，则进入配对交易策略. 更为有趣的是，应用这个方法可以连续监

测价格过程并且发出信号，如果序列接受平稳性假设，则自动进入一个配对交易策略. 在 9.4.4 中将详细介绍这种方法. ◀

例 9.4.1 的情形在数学上称为协整，在 9.4.2 中将给出它的严格定义.

9.4.1 平稳 AR(1)模型的最小二乘估计量

在模型(9.21)中，扰动项 u_t 是独立同分布的. 为了得到是接受单位根假设或接受平稳性假设的推理过程，现在将模型(9. 21)推广到有鞅差分序列误差项的情形. 设

$$Y_t = \rho Y_{t-1} + u_t, \qquad t = 1,2,\cdots, \tag{9.22}$$

其中$\{u_t\}$是关于滤子 $\mathcal{F}_t=\sigma(u_s: s\leqslant t)$的一个鞅差分序列，即

$$E(u_t \mid \mathcal{F}_{t-1}) = 0, \quad t = 1,2,\cdots.$$

除此之外，还假设存在有限常数 σ^2，γ_4 使得

364

$$E(u_t^2 \mid \mathcal{F}_{t-1}) = \sigma^2 \quad \text{和} \quad E(u_t^4 \mid \mathcal{F}_{t-1}) = \gamma_4 \text{(a. s.)}.$$

这些假设使得我们能够运用第 3 章讨论过的鞅差分序列和鞅差分阵列的极限定理.

给定 Y_1，…，Y_T，ρ 的最小二乘估计量$\hat{\rho}_T$ 为

$$\hat{\rho}_T = \underset{\rho\in \mathbf{R}}{\operatorname{argmin}} \sum_{t=1}^{T} (Y_t - \rho Y_{t-1})^2.$$

通过直接计算得到

$$\hat{\rho}_T = \frac{\sum_{t=1}^{T} Y_t Y_{t-1}}{\sum_{t=1}^{T} Y_{t-1}^2},$$

它与误差项是独立同分布的正态分布时的最大似然估计是一致的，在 i. i. d. 正态误差项时，分母 $\sum_t Y_{t-1}^2$得到 Fisher 信息量.

接下来讨论当$|\rho|<1$ 时$\hat{\rho}_T$ 的渐近分布. 下面的定理表明$\hat{\rho}_T$ 以收敛速度 $1/\sqrt{T}$依概率收敛于ρ，且是渐近正态的. 作为 8.2 中讨论过的多元回归模型最小二乘估计的一个特例，这里给出一个直接的证明来突出下一小节中介绍的单整时间序列对应的一些统计量的完全不同性质.

定理 9.4.2 设$\{u_t\}$是一个严平稳的鞅差分序列，且 $E(u_t^2 \mid \mathcal{F}_{t-1})=\sigma^2$，$E(u_t^4 \mid \mathcal{F}_{t-1})=\gamma_4<\infty$. 则当 $T\to\infty$时，有

(ⅰ)

$$\frac{1}{T}\sum_{t=1}^{T} Y_{t-1}^2 \xrightarrow{P} \frac{\sigma^2}{1-\rho^2}, \tag{9.23}$$

(ⅱ)

$$\frac{1}{\sqrt{T}}\sum_{t=1}^{T} Y_{t-1} u_t \xrightarrow{d} N(0,\sigma^4/(1-\rho^2)), \tag{9.24}$$

从而得到

$$\sqrt{T}(\hat{\rho}_T-\rho)\xrightarrow{d}N(0,(1-\rho^2)),\quad T\to\infty,$$

其中 $\sigma^2=E(\varepsilon_1^2)$.

证明 注意到

$$\hat{\rho}_T-\rho=\frac{\sum_{t=1}^{T}(Y_tY_{t-1}-\rho Y_{t-1}^2)}{\sum_{t=1}^{T}Y_{t-1}^2}=\frac{\sum_{t=1}^{T}\varepsilon_tY_{t-1}}{\sum_{t=1}^{T}Y_{t-1}^2},$$

所以

$$\sqrt{T}(\hat{\rho}_T-\rho)=\frac{T^{-1/2}\sum_{t=1}^{T}Y_{t-1}u_t}{T^{-1}\sum_{t=1}^{T}Y_{t-1}^2}.$$

先看分母. 结论(ⅰ)由命题 8.4.1 可得，因为 $E(u_t^2\mid\mathcal{F}_{t-1})=\sigma^2$，$E(u_t^4\mid\mathcal{F}_{t-1})=\gamma_4<\infty$. 接下来，记

$$\xi_t=Y_{t-1}u_t,\quad t=1,2,\cdots,$$

下面证明$\{\xi_t\}$是关于滤子 $\mathcal{F}_t=\sigma(u_s:\ s\leqslant t)$的一个严平稳鞅差分序列. 因为 AR(1)方程的平稳解为下列线性过程

$$Y_t=\sum_{i=0}^{\infty}\rho^iu_{t-i}.$$

由此可得 $\xi_t=Y_{t-1}u_t$ 具有 $f(u_t,\ u_{t-1},\ \cdots)$的形式，其中 f：$\mathbf{R}^\infty\to\mathbf{R}$ 是 Borel 可测函数. 所以 ξ_t 是严平稳的. 又因为 $E(Y_{t-1}^2)=\sigma^2\sum_{i=0}^{\infty}\rho^{2i}<\infty$，所以

$$E|\xi_t|\leqslant\sqrt{E(Y_{t-1}^2)E(u_t^2)}<\infty,\text{对所有的 } t.$$

进一步对所有的 t,

$$E(\xi_t\mid\mathcal{F}_{t-1})=E(Y_{t-1}u_t\mid Y_{t-1})=Y_{t-1}E(u_t\mid\mathcal{F}_{t-1})=0,(\text{a.s.}).$$

由此可得$\{\xi_t\}$是一个严平稳的鞅差分序列. 为了证明$\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\xi_t$ 满足中心极限定理，对 $\varepsilon>0$，考虑 Lindeberg 条件

$$L_T(\varepsilon)=\frac{1}{T}\sum_{t=1}^{T}E(E(Y_{t-1}^2u_t^2\mathbf{1}(Y_{t-1}u_t>\varepsilon)\mid\mathcal{F}_{t-1})),$$

它如下形式

$$\frac{1}{T}\sum_{t=1}^{T}\int y^2\int u_t^2(\omega)\mathbf{1}(|yu_t(\omega)|>\sqrt{T}\varepsilon)\mathrm{d}P_{u_t|Y_{t-1}=y}(\omega)\mathrm{d}P_{Y_{t-1}}(y).$$

显然，对任意固定的 ω 和 y，当 $T\to\infty$时，示性函数收敛于 0. 另外因为被积函数是有界的(其界为可积随机变量 u_t^2)，这表明 $E(u_T^2\mid\mathcal{F}_{t-1})<\infty$，a. s.，由条件控制收敛定理知：对所有固定的 t，当 $T\to\infty$时

$$\int u_t^2(\omega)\mathbf{1}(|yu_t(\omega)| > \sqrt{T}\varepsilon)\mathrm{d}P_{u_t|Y_{t-1}=y}(\omega) \to 0,$$

再一次应用控制收敛定理得：对每个 t，当 $T\to\infty$ 时

$$E(Y_{t-1}^2 u_t^2 \mathbf{1}(Y_{t-1}u_t > \sqrt{T}\varepsilon)) \to 0,$$

而由严平稳性，这些被加项不依赖于 t，所以，

366

$$L_T(\varepsilon) \to 0, \quad T \to \infty.$$

最后，讨论

$$\frac{1}{T}\sum_{t=1}^{T} E(Y_{t-1}^2 u_t^2 \mid \mathcal{F}_{t-1})$$

的收敛性. 因为 $E(u_t^2 \mid \mathcal{F}_{t-1})=\sigma^2$，所以，当 $T\to\infty$ 时

$$\frac{1}{T}\sum_{t=1}^{T} E(Y_{t-1}^2 u_t^2 \mid \mathcal{F}_{t-1}) = \frac{1}{T}\sum_{t=1}^{T} Y_{t-1}^2 E(u_t^2 \mid \mathcal{F}_{t-1}) \to \frac{\sigma^4}{1-\rho^2},$$

所以定理 B. 7. 2 的条件(ⅰ)和(ⅱ)满足，从而得到：当 $T\to\infty$ 时

$$\frac{1}{\sqrt{T}}\sum_{t=1}^{T} Y_{t-1}u_t \xrightarrow{d} Z \sim N(0,\sigma^4/(1-\rho^2)),$$

这就证明了结论(ⅱ). 把这一结论与(9. 23)式结合在一起，并应用 Slutzky 引理，得到

$$\sqrt{T}(\hat{\rho}_T-\rho) = \frac{T^{-1/2}\sum_{t=1}^{T} Y_{t-1}u_t}{T^{-1}\sum_{t=1}^{T} Y_{t-1}^2} \xrightarrow{d} \left(\frac{\sigma^2}{1-\rho^2}\right)^{-1} Z \sim N(0,(1-\rho^2)), \quad T\to\infty,$$

证毕. ■

上面的结果使得人们能够检验关于 AR 参数 ρ 的统计假设，以及建立该参数的置信区间.

设检验问题为

$$H_0:\rho=\rho_0 \quad \longleftrightarrow \quad H_1:\rho\neq\rho_0,$$

其中，$\rho_0\in(-1,\ 1]$是某个已知的数. 这类问题可以按下列方法检验：对于一个给定的置信水平 $\alpha\in(0,\ 1)$，若 $|S_T|>z_{1-\alpha/2}$，则拒绝原假设，其中

$$S_T = \sqrt{T}\,\frac{\hat{\rho}_T-\rho_0}{\sqrt{1-\hat{\rho}_T^2}},$$

且 $z_{1-\alpha/2}$ 是 $N(0,\ 1)$分布的$(1-\alpha/2)$-分位数. 渐近收敛概率为 $1-\alpha$ 的置信区间为

$$\left[\hat{\rho}_T - z_{1-\alpha/2}\,\frac{\sqrt{1-\hat{\rho}_T^2}}{\sqrt{T}},\hat{\rho}_T + z_{1-\alpha/2}\,\frac{\sqrt{1-\hat{\rho}_T^2}}{\sqrt{T}}\right].$$

注意，估计量 $\hat{\rho}_T$ 的渐近分布不依赖于测量的尺度. 然而，它需要估计扰动项 u_t 的离差 σ. 一个自然的方法是：记残差为

367

$$\hat{u}_t = Y_t - \hat{\rho}_T Y_{t-1}, \quad t=1,\cdots,T.$$

σ^2 的估计量为

$$\hat{\sigma}_T^2 = \frac{1}{T}\sum_{t=1}^{T}\hat{u}_t^2.$$

该估计量是一致估计量，其证明是直接的.

命题 9.4.3 在定理 9.4.2 的条件下，

$$\hat{\sigma}_T^2 \xrightarrow{P} \sigma^2, \quad T \to \infty.$$

证明 由 $\hat{u}_t = Y_t - \hat{\rho}_T Y_{t-1} = (\rho - \hat{\rho}_T)Y_{t-1} + u_t$ 得到

$$\hat{u}_t^2 = (\rho - \hat{\rho}_T)^2 Y_{t-1}^2 + 2(\rho - \hat{\rho}_T)Y_{t-1}u_t + u_t^2.$$

由强大数定律知上式最后一项的算术平均 a.s. 收敛于 σ^2. 如能证明当 $T\to\infty$ 时

$$A_T = (\rho - \hat{\rho}_T)^2 \frac{1}{T}\sum_{t=1}^{T} Y_{t-1}^2 \xrightarrow{P} 0, \quad B_T = 2(\rho - \hat{\rho}_T)\frac{1}{T}\sum_{t=1}^{T} Y_{t-1}u_t \xrightarrow{P} 0,$$

则结论得证. 下面考虑 A_T，由 $\sqrt{T}(\hat{\rho}_T - \rho)$ 的渐近正态性知 $\hat{\rho}_T - \rho \xrightarrow{P} 0$，从而 $(\hat{\rho}_T - \rho)^2 \xrightarrow{P} 0$. 进一步，在定理 9.4.2 的证明中已经得到 $T^{-1}\sum_{t=1}^{T} Y_{t-1}^2$ 依概率收敛于一个有限常数. 因此，当 $T\to\infty$ 时，$A_T \xrightarrow{P} 0$. $B_T \xrightarrow{P} 0$ 类似可证. ■

9.4.2 整合度的非参数定义

$|\rho|<1$ 时的 AR(1) 方程(9.22)的平稳解 Y_t 可以表示为一个线性过程 $Y_t = \sum_{i=0}^{\infty}\rho^i u_{t-i}$，因此满足泛函中心极限定理. 从而部分和过程

$$S_T(u) = \frac{1}{\sqrt{T}}\sum_{t=1}^{\lfloor Tu \rfloor} Y_t, \quad u \in [0,1], \quad T \geqslant 1$$

满足

$$S_T(u) \Rightarrow \eta B(u), \quad T \to \infty, \tag{9.25}$$

其中 B 是一个标准布朗运动且 $\eta = \sqrt{\eta^2}$，这里

$$\eta^2 = EY_0^2 + 2\sum_{k=0}^{\infty} E(Y_0 Y_k)$$

是 $\{Y_n\}$ 的**长期方差**. η^2 的一致估计已经在 8.8 节讨论过了，见定理 8.8.6. 368

如果 $\rho=1$，则 Y_t 是一个随机游动，其增量是鞅差分序列，因而满足

$$T^{-1/2} Y_{\lfloor Ts \rfloor} \Rightarrow \sigma B(s), \quad T \to \infty, \tag{9.26}$$

其中 $Y_{\lfloor Ts \rfloor}$，$s\in[0, 1]$ 是 Y_1，…，Y_T 的**标准过程**，即

$$Y_{\lfloor Ts \rfloor} = Y_0 + \sum_{i=0}^{\lfloor Ts \rfloor} \Delta Y_t.$$

且一阶差分

$$\Delta Y_t = Y_t - Y_{t-1}$$

是平稳的且满足泛函中心极限定理. 如果随机游动 Y_t 及 Y_t 的增量分别是相关的时间序列，FCLT(9.25)和弱收敛(9.25)仍成立.

对一个给定的统计量，比如 U_T，它通过部分和过程 S_T 依赖于时间序列数据 $\{Y_t\}$，即

$$U_T = U(S_T),$$

其中，S_T 所有的轨道在 $\mathcal{D}$ 中，U：$\mathcal{D}\to\mathbf{R}$ 是一个定义在 $\mathcal{D}$ 上的函数. 如果映射 U 是光滑的(即当 $T\to\infty$ 时，如果 S_T 弱收敛于 ηB，那么 $U(S_T)$ 依分布收敛于 $U(\eta B)$)，则对任意满足 FCLT(9.25)的时间序列 $\{Y_t\}$，统计量 U_T 的渐近性质是已知的，这就得到了一大类时间序列的一般极限定理. 类似地，根据 FCLT(9.26)，任一标准过程的光滑函数构成的统计量(即 $V_T=V(T^{-1/2}Y_{\lfloor T\cdot\rfloor})$)的渐近分布由随机元 $V(\sigma B)$ 给出. 对于各种单位根统计量，都会遇到这种情形，这就导出了下列非参数定义.

定义 9.4.4(单整时间序列)

(ⅰ) 称一个离散时间的随机过程 $\{Y_t\}$ 为 **0 阶单整的**，记为 $Y_t\sim I(0)$，如果它满足一个泛函中心极限定理，即存在常数 $\eta\in(0,\infty)$，使得当 $T\to\infty$ 时

$$\frac{1}{\sqrt{T}}\sum_{t=1}^{\lfloor Ts\rfloor}Y_t\Rightarrow\eta B(s),$$

其中，B 是一个标准布朗运动.

(ⅱ) 称一个离散时间的随机过程 $\{Y_t\}$ 为 **1 阶单整的**，记为 $Y_t\sim I(1)$，如果它对应的标准过程满足

369

$$T^{-1/2}Y_{\lfloor Ts\rfloor}\Rightarrow\sigma B(s),\quad T\to\infty,$$

并且它的一阶差分 ΔY_t 是满足泛函中心极限定理的弱平稳序列. ■

评注 9.4.5 需要注意的是在上面的定义中，$I(0)$ 包含了非平稳过程，只要它们满足带布朗运动的泛函中心极限定理. 然而，计量经济学家研究的模型通常得到的是平稳 $I(0)$ 序列. ■

上面的定义引出下列协整时间序列的定义.

定义 9.4.6 称二维时间序列 $\{(Y_t,X_t):t=1,2,\cdots\}$，其中 $Y_t\sim I(1)$，$X_t\sim I(1)$，**为协整的**，如果存在一个常数 $a\neq0$，使得 $Y_t-aX_t\sim I(0)$. 一般地，称一个 d 维时间序列 $\{X_t\}$，$X_t=(X_{t1},\cdots,X_{td})$，其中，$X_{ti}\sim I(1)(i=1,\cdots,d)$ 为协整的，如果存在某个非零向量 $a\in\mathbf{R}^d$，使得 $X_t\sim I(0)$. ■

9.4.3 Dickey-Fuller 检验

仍然研究模型(9.22). 检验有单位根的原假设 $H_0:\rho=0$，平稳的备择假设 $H_1:|\rho|<1$ 的一个广泛使用的统计量是 Dickey-Fuller 统计量

$$D_T = T(\hat{\rho}_T-1).$$

下列定理给出了在有单位根的原假设下$\left(\text{即 } Y_t=\sum_{i=1}^{t}\varepsilon_i \text{ 是 } I(1)\text{的}\right)$，一些渐近分布结果. 它表明 $\hat{\rho}_T$ 的收敛速度是 T，即在有单位根的情形下，$\hat{\rho}_T$ 收敛于真实值 $\rho_0=1$ 的速度快于平稳条件下的收敛速度 $\sqrt{T}$. 这意味着，在随机游动的序列中，$\hat{\rho}_T$ 是超一致估计量.

定理 9.4.7 设部分和过程 $S_T(u)=T^{-1/2}\sum_{i=1}^{\lfloor Tu\rfloor}\varepsilon_i$，$u\in[0,1]$满足

$$S_T \Rightarrow \eta B, \quad T\to\infty,$$

其中，$\eta\in(0,\infty)$是一个常数，且

$$\frac{1}{T}\sum_{t=1}^{T}\varepsilon_i^2 \xrightarrow{P} \sigma^2, \quad T\to\infty.$$

那么

（i）$\frac{1}{T}\sum_{t=1}^{T} Y_{t-1}\varepsilon_t \xrightarrow{d} \frac{\eta^2}{2}\left(B^2(1)-\frac{\sigma^2}{\eta^2}\right)$，当 $T\to\infty$ 时，

（ii）$\frac{1}{T^2}\sum_{t=1}^{T} Y_{t-1}^2 \xrightarrow{d} \eta^2\int_0^1 B^2(r)\,\mathrm{d}r$，当 $T\to\infty$ 时，

370

且这两个收敛是联合收敛的．进一步，最小二乘估计量$\hat{\rho}_T$满足

$$T(\hat{\rho}_T-1)\xrightarrow{d}\frac{1}{2}\frac{B^2(1)-\frac{\sigma^2}{\eta^2}}{\int_0^1 B^2(r)\,\mathrm{d}r}, \quad T\to\infty.$$

证明 令 $S(T)=\sqrt{T}S_T(1)=\sum_{i=1}^{T}\varepsilon_i$．首先证明

$$\sum_{t=1}^{T}Y_{t-1}\varepsilon_t=\frac{1}{2}\left(S(T)^2-\sum_{t=1}^{T}\varepsilon_t^2\right). \tag{9.27}$$

注意到

$$\sum_{t=1}^{T}Y_{t-1}\varepsilon_t=\sum_{t=1}^{T}\left(\sum_{j=1}^{t-1}\varepsilon_j\right)\varepsilon_t=\sum_{1\leqslant i<j\leqslant T}\varepsilon_i\varepsilon_j$$

是对称矩阵$(\varepsilon_i\varepsilon_j)_{i,j}$主对角线元素的和．因此

$$\left(\sum_{t=1}^{T}\varepsilon_t\right)^2=\sum_{i,j=1}^{T}\varepsilon_i\varepsilon_j=2\sum_{i<j}\varepsilon_i\varepsilon_j+\sum_{t=1}^{T}\varepsilon_t^2,$$

这就得到

$$\sum_{t=1}^{T}Y_{t-1}\varepsilon_t=\sum_{i<j}\varepsilon_i\varepsilon_j=\frac{1}{2}\left[\left(\sum_{t=1}^{T}\varepsilon_t\right)^2-\sum_{t=1}^{T}\varepsilon_t^2\right]=\frac{1}{2}\left[S(T)^2-\sum_{t=1}^{T}\varepsilon_t^2\right],$$

其中 $S(T)^2=TS_T^2(1)$满足

$$\frac{1}{T}\sum_{t=1}^{T}Y_{t-1}\varepsilon_t=\frac{1}{2}\left[S_T^2(1)-\frac{1}{T}\sum_{t=1}^{T}\varepsilon_t^2\right]. \tag{9.28}$$

现在考虑

$$\frac{1}{T^2}\sum_{t=1}^{T}Y_{t-1}^2=\frac{1}{T}\sum_{t=1}^{T}\frac{1}{T}\left(\sum_{t=1}^{t-1}\varepsilon_i\right)^2.$$

注意到 $1/T=\int_0^1 \mathbf{1}_{[(t-1)/T,t/T]}(r)\,\mathrm{d}r$，由向下取整函数的定义知

$$\sum_{i=1}^{\lfloor Tr\rfloor}\varepsilon_i=\sum_{i=1}^{t-1}\varepsilon_i, \quad \text{对 } t-1\leqslant Tr<t\Leftrightarrow r\in[(t-1)/T,t/T), \quad t=1,\cdots,T.$$

我们将 1 划分为 $1_{[0,1]}=\sum_{t=1}^{T}\mathbf{1}_{[(t-1)/T,t/T]}$，并由

$$\frac{1}{T}\sum_{t=1}^{t-1}\varepsilon_i = \sum_{i=1}^{t-1}\int_{\frac{t-1}{T}}^{\frac{t}{T}} dr = \int_{\frac{t-1}{T}}^{\frac{t}{T}}\sum_{i=1}^{\lfloor Tr \rfloor}\varepsilon_i dr$$

371

得到

$$\begin{aligned}\frac{1}{T^2}\sum_{t=1}^{T}Y_{t-1}^2 &= \frac{1}{T}\sum_{t=1}^{T}\int_0^1\Big(\sum_{i=1}^{\lfloor Tr \rfloor}\varepsilon_i\Big)^2\mathbf{1}_{[(t-1)/T,t/T)}(r)dr \\ &= \sum_{t=1}^{T}\int_0^1 S_T^2(r)\mathbf{1}_{[(t-1)/T,t/T)}(r)dr \\ &= \int_0^1 S_T^2(r)\sum_{t=1}^{T}\mathbf{1}_{[(t-1)/T,t/T)}(r)dr.\end{aligned}$$

因此

$$\frac{1}{T^2}\sum_{t=1}^{T}Y_{t-1}^2 = \int_0^{1-} S_T^2(r)dr. \tag{9.29}$$

为了证明联合弱收敛，考虑二维随机向量$\Big(T^{-1}\sum\limits_{t=1}^{T}Y_{t-1}\varepsilon_t,\ T^{-2}\sum\limits_{t=1}^{T}\varepsilon_t^2\Big)$. 结合(9.28)式和(9.29)式，得到

$$\begin{aligned}\Big(\frac{1}{T}\sum_{t=1}^{T}Y_{t-1}\varepsilon_t,\frac{1}{T^2}\sum_{t=1}^{T}\varepsilon_t^2\Big) &= \Big(\frac{1}{2}\Big[S_T^2(1)-\frac{1}{T}\sum_{t=1}^{T}\varepsilon_t^2\Big],\int_0^{1-}S_T^2(r)dr\Big) \\ &= \Big(\frac{1}{2}[S_T^2(1)-\sigma^2],\int_0^{1-}S_T^2(r)dr\Big)+o_P(1) \\ &= \phi(S_T)+o_P(1),\end{aligned}$$

其中函数 ϕ：$D([0,\ 1];\ \mathbf{R})\to\mathbf{R}^2$ 定义为

$$\phi(f) = \Big(\frac{1}{2}[f^2(1)-\sigma^2],\int_0^{1-}f^2(r)dr\Big),\quad f\in D([0,1];\mathbf{R}).$$

ϕ 是连续的，见例 B. 2. 2. 因此，由连续映射定理

$$\phi(S_T)\Rightarrow\phi(\eta B) = \Big(\frac{\eta^2}{2}\Big[B^2(1)-\frac{\sigma^2}{\eta^2}\Big],\eta^2\int_0^1 B^2(r)dr\Big),\quad T\to\infty$$

就证明了结论(ⅰ)和(ⅱ). 因为函数 $F(x,\ y)=x/y$ 在 $\mathbf{R}\times\mathbf{R}-\{0\}$ 上是连续的，且 $\int_0^1 B^2(r)dr$ (a. s.)是正的，由连续映射定理可得 $D_T=T(\hat{\rho}_T-1)$的渐近性.

将上述定理的结论(ⅰ)和(ⅱ)与给出平稳 AR(1)模型对应结论的定理 9.4.2 进行比较. 对 $I(1)$序列，统计量 $\sum_t Y_{t-1}^2$，$\sum_t Y_{t-1}\varepsilon_t$(适当缩放后)的渐近性质，比如收敛速度和极限分布的类型，是完全不同的.

372

定理 9.4.7(ⅰ)中的极限过程是 $B^2(1)\sim\chi^2(1)$的线性函数，可以通过下述统计量

$$\frac{1}{T}\sum_{t=1}^{T}Y_{t-1}\varepsilon_t = \Big(\frac{1}{T}\sum_t Y_{t-1}^2\Big)(\hat{\rho}_T-1)$$

的变换来得到一个 χ^2-分布的随机变量，这就是 Phillips-Durlanf 统计量

$$\frac{2}{\eta^2}\left(\frac{1}{T}\sum_{t=1}^{T}Y_{t-1}^2\right)(\hat{\rho}_T-1)+\frac{\sigma^2}{\eta^2}\Rightarrow B^2(1)-\chi^2(1), \quad T\to\infty$$

(在单位根原假设下). 在这个公式中，可以用一致估计量来代替未知的参数 η^2，σ^2.

9.4.4 检测单位根和平稳性

现在的问题是如何从一个时间序列中检测出定义 9.4.4 中的 $I(0)$平稳到 $I(1)$非平稳的变点. 记真实的确定性未知变点为 q，并设它与时间 T 成正比，对某个 $\vartheta\in(0, 1)$

$$q=\lfloor T\vartheta\rfloor.$$

下列两种变点模型是人们感兴趣的.

I(1)－I(0)变点：假设变点前序列的观测值为

$$\{Y_0,\cdots,Y_{\lfloor T\vartheta\rfloor-1}\}\sim I(1),$$

它是定义 9.4.4(ⅱ)意义下的 $I(1)$序列，而观测值

$$\{Y_{\lfloor T\vartheta\rfloor},\cdots,Y_T\}\sim I(0)$$

是定义 9.4.4(ⅰ)意义下的 $I(0)$序列.

I(0)－I(1)变点：这里，序列的第一部分是 $I(0)$序列

$$\{Y_0,\cdots,Y_{\lfloor T\vartheta\rfloor-1}\}\sim I(0),$$

从变点 q 开始变为 $I(1)$序列

$$\{Y_{\lfloor T\vartheta\rfloor},\cdots,Y_T\}\sim I(1).$$

I(0)检测：检测从 $I(1)$到 $I(0)$的变点，可以使用序贯 KPSS 或 Dickey-Fuller 检验统计量. 使用 KPSS 检验方法如下：如果 $s\in[0, 1/T]$，定义 $U_T(s)=0$，且

$$U_T(s)=\frac{\lfloor Ts\rfloor^{-3}\sum_{i=1}^{\lfloor Ts\rfloor}\left(\sum_{j=1}^{i}Y_j\right)^2K_h(i-\lfloor Ts\rfloor)}{\lfloor Ts\rfloor^{-2}\sum_{j=1}^{\lfloor Ts\rfloor}Y_j^2}, s\in[1/T,1]. \tag{9.30}$$

称 U_T 为一个**序贯 KPSS 过程**，其中，K 是一个零均值，有有限非零二阶矩的核，$K_h(\cdot)=K(\cdot/h)/h$ 是 K 的缩放形式. 它用于确保较远的时间点 i 对应的部分和 $\sum_{j=1}^{i}Y_j$ 对统计量的影响小于当前时刻 $t=\lfloor Ts\rfloor$ 附近的时间点 i 对应的部分和对统计量的影响. 若核 K 的支集为$[-1, 1]$，只有位于时间窗口$[t-h, t]$的观测值是有用的. 因此带宽 h 确定了样本的范围. 单整阶的推断需要相当大的有效样本数，并且不假设观测到的时间序列 Y_1，Y_2，…趋于0. 因此，与 8.5 节中讨论的非参数平滑过程不同，这里不假设当样本数趋于∞时带宽趋于 0，而是假设 $h=h_T$ 是 T 的一个函数，使之满足

373

$$\lim_{T\to\infty}\frac{T}{h_T}=\zeta\in[1,\infty),$$

这个条件确保了当 T 无限增大时，能在过程中被有效地使用的观测值的数目趋于∞. 在实践中，人们是先固定 ζ，然后选择 $h_T=[T/\zeta]$.

如果 $U_T(t/T)$首次落在一个固定值之下，停止监测并发出从 $I(1)$到 $I(0)$变点的信号.

也就是说，考虑停时

$$R_T = R_T(c) = \min(k \leqslant t \leqslant T : U_T(t/T) < c),$$

其中 c 是某个极限常数，它确保在不超过 c 的范围内统计量的性质成立，k 为检测开始时刻．对于下面将要讨论的渐近结果，设

$$k = \lfloor \kappa T \rfloor, \quad \kappa \in (0,1). \tag{9.31}$$

前 $k-1$ 个观测值可以作为研究与监测过程无关的其他有趣的量的样本．

确定极限常数 c 的自然方法是确保过程停止的 I 类差错率

$$\alpha_T = P_0(R_T \leqslant T)$$

收敛于某个给定的名义显著水平 $\alpha \in (0, 1)$

$$\lim_{T \to \infty} \alpha_T = \alpha,$$

其中 P_0 表示在原假设(过程 $\{Y_t\}$ 是 $I(1)$)下求概率．

下列定理给出了序贯 KPSS 过程在原假设下的分布．

定理 9.4.8 设 $\{Y_t\}$ 是定义 9.4.4 中定义的 $I(1)$ 过程，则序贯 KPSS 过程是弱收敛的，即当 $T \to \infty$ 时，在 $D[\kappa, 1]$ 上，

374

$$U_T(s) \Rightarrow \mathcal{U}_1(s) = \frac{\zeta s^{-1} \int_0^s K(\zeta(r-s)) \left[\int_0^r B(t) \mathrm{d}t \right]^2 \mathrm{d}r}{\int_0^s B(r)^2 \mathrm{d}r}. \tag{9.32}$$

证明 注意到

$$U_T(s) = \frac{\left(\frac{T}{\lfloor Ts \rfloor} \right)^3 X_T(s)}{\left(\frac{T}{\lfloor Ts \rfloor} \right)^2 Z_T(s)},$$

如果定义

$$X_T(s) = \frac{1}{T^3} \sum_{i=1}^{\lfloor Ts \rfloor} \left(\sum_{j=1}^{i} Y_j \right)^2 K_h(i - \lfloor Ts \rfloor),$$

$$Z_T(s) = \frac{1}{T^2} \sum_{i=1}^{\lfloor Ts \rfloor} Y_i^2 = \int_0^{\lfloor Ts \rfloor / T} \left(\frac{1}{\sqrt{T}} Y_{\lfloor Tr \rfloor} \right)^2 \mathrm{d}r,$$

其中

$$S_T(s) = \frac{1}{\sqrt{T}} \sum_{t=1}^{\lfloor Ts \rfloor} Y_t, \quad s \in [0,1].$$

如果能够证明：当 $T \to \infty$ 时

$$\frac{X_T(s)}{Z_T(s)} \Rightarrow s\, \mathcal{U}_1(s), \tag{9.33}$$

则由 Slutzky 引理及引理 B.2.4 本定理得证．为了证明(9.33)式，用一个布朗运动的右可积泛函以及在定理条件下部分和过程 S_T 的极限过程，来近似分子分母．注意，为了记号简洁，当涉及一个等价的新概率空间时，将不改变记号．由 Skorohod 表示定理，不失一般性，设当 $T \to \infty$ 时，

$$\sup_{u\in[0,1]}\left|\frac{Y_{\lfloor Tu\rfloor}}{\sqrt{T}}-\sigma B(u)\right|\to 0,\quad \text{a. s.}. \tag{9.34}$$

为了叙述简单，当在新概率空间上处理等价形式时，不改变记号．将 $X_T(s)$表示为一个迭代黎曼积分．由

$$t=\lfloor Tr\rfloor \Leftrightarrow r\in\left[\frac{t}{T},\frac{t+1}{T}\right)$$

得到

$$\frac{1}{T}\Big(\frac{1}{T}\sum_{j=1}^{t}\frac{Y_j}{\sqrt{T}}\Big)^2K\Big(\frac{t-\lfloor Ts\rfloor}{h}\Big)$$
$$=\int\mathbf{1}_{\left[\frac{t}{T},\frac{t+1}{T}\right)}(r)\Big(\frac{1}{T}\sum_{j=1}^{\lfloor Tr\rfloor}\frac{Y_j}{\sqrt{T}}\Big)^2K\Big(\frac{\lfloor Tr\rfloor-\lfloor Ts\rfloor}{h}\Big)\mathrm{d}r.$$

同样地，

375

$$j=\lfloor Tu\rfloor\Leftrightarrow u\in\left[\frac{j}{T},\frac{j+1}{T}\right),$$

得到

$$\frac{1}{T}\sum_{j=1}^{\lfloor Tr\rfloor}\frac{Y_j}{\sqrt{T}}=\sum_{j=1}^{\lfloor Tr\rfloor}\int\mathbf{1}_{\left[\frac{j}{T},\frac{j+1}{T}\right)}(u)\frac{Y_{\lfloor Tu\rfloor}}{\sqrt{T}}\mathrm{d}u.$$

综上所述，得到

$$\begin{aligned}X_T(s)&=\frac{T}{h}\frac{1}{T}\sum_{t=1}^{T}\mathbf{1}(t\leqslant\lfloor Ts\rfloor)\Big(\frac{1}{T}\sum_{j=1}^{t}\frac{Y_j}{\sqrt{T}}\Big)^2K\Big(\frac{t-\lfloor Ts\rfloor}{h}\Big)\\&=\frac{T}{h}\sum_{t=1}^{T}\mathbf{1}(t\leqslant\lfloor Ts\rfloor)\int_{\frac{t}{T}}^{\frac{t+1}{T}}\Big(\sum_{j=1}^{\lfloor Tr\rfloor}\int_{\frac{j}{T}}^{\frac{j+1}{T}}\frac{Y_{\lfloor Tu\rfloor}}{\sqrt{T}}\mathrm{d}u\Big)^2K\Big(\frac{\lfloor Tr\rfloor-\lfloor Ts\rfloor}{h}\Big)\mathrm{d}r\\&=\frac{T}{h}\int_{\frac{1}{T}}^{\frac{\lfloor T(s+1/T)\rfloor}{T}}\Big(\int_{\frac{1}{T}}^{\frac{\lfloor T(r+1/T)\rfloor}{T}}\frac{Y_{\lfloor Tu\rfloor}}{\sqrt{T}}\mathrm{d}u\Big)^2K\Big(\frac{\lfloor Tr\rfloor-\lfloor Ts\rfloor}{h}\Big)\mathrm{d}r,\end{aligned}$$

其中，用到$\frac{\lfloor Ts\rfloor+1}{T}=\frac{\lfloor T(s+1)/T\rfloor}{T}$这一事实．注意到由(9.34)式可推出

$$\int_a^b\frac{Y_{\lfloor Tu\rfloor}}{\sqrt{T}}\mathrm{d}u\Rightarrow\int_a^b\sigma B(u)\mathrm{d}u,$$

对任何 $0\leqslant a\leqslant b\leqslant 1$ 成立，且当 $T\to\infty$时，

$$I_T(B)=\sup_{a,b\in[0,1],a\leqslant b}\int_a^b|\sigma B(u)|\mathrm{d}u=O_P(1),$$

进一步，当 $T\to\infty$时有

$$\begin{aligned}&\sup_{r\in[\kappa,1]}\left|\Big(\int_{\frac{1}{T}}^{\frac{\lfloor T(r+1/T)\rfloor}{T}}\frac{Y_{\lfloor Tu\rfloor}}{\sqrt{T}}\Big)^2-\Big(\int_{\frac{1}{T}}^{\frac{\lfloor T(r+1/T)\rfloor}{T}}\sigma B(u)\Big)^2\right|\\&\leqslant\sup_{r\in[\kappa,1]}\Big(\int_0^1|\sigma B(u)|\mathrm{d}u+\int_0^1\left|\frac{Y_{\lfloor Tu\rfloor}}{\sqrt{T}}\right|\mathrm{d}u\Big)\int_0^1\left|\frac{Y_{\lfloor Tu\rfloor}}{\sqrt{T}}-\sigma B(u)\right|\mathrm{d}u\\&\leqslant\sup_{r\in[\kappa,1]}\Big(\int_0^1|\sigma B(u)|\mathrm{d}u+\int_0^1\left|\frac{Y_{\lfloor Tu\rfloor}}{\sqrt{T}}\right|\mathrm{d}u\Big)\sup_{u\in[0,1]}\left|\frac{Y_{\lfloor Tu\rfloor}}{\sqrt{T}}-\sigma B(u)\right|\end{aligned}$$

376

$$= o_P(1),$$

因为括号中的表达式是 $O_P(1)$. 再结合 K, T/h 的有界性，从而得到：当 $T\to\infty$时

$$X_T(s) = \frac{T}{h}\int_{\frac{1}{T}}^{\frac{\lfloor T(s+1/T)\rfloor}{T}}\left(\int_{\frac{1}{T}}^{\frac{\lfloor T(r+1/T)\rfloor}{T}}\sigma B(u)\mathrm{d}u\right)^2 K\left(\frac{\lfloor Tr\rfloor-\lfloor Ts\rfloor}{h}\right)\mathrm{d}r + o_P(1),$$

其中 $o_P(1)$项在 $s\in[\kappa,1]$时是一致的. 由 $T/h\to\zeta$ 和 K 的 Lipschitz 连续性，进一步得到：当 $T\to\infty$时，

$$X_T(s) = \zeta\int_{\frac{1}{T}}^{\frac{\lfloor T(s+1/T)\rfloor}{T}}\left(\int_{\frac{1}{T}}^{\frac{\lfloor T(r+1/T)\rfloor}{T}}\sigma B(u)\mathrm{d}u\right)^2 K(\zeta(r-s))\mathrm{d}r + o_P(1),$$

其中 $o_P(1)$项在 $s\in[\kappa,1]$时也是一致的. 分别用$(0,s]$，$(0,r]$代替积分区间，得到误差项在$(0,s]$，$(0,r]$中一致有阶 $o_P(1)$. 例如，当 $T\to\infty$时

$$\begin{aligned}
&\sup_{r\in[\kappa,1]}\left|\int_{\frac{1}{T}}^{\frac{\lfloor T(r+1/T)\rfloor}{T}}\sigma B(u)\mathrm{d}u - \int_0^r\sigma B(u)\mathrm{d}u\right| \\
&\leqslant \frac{1}{T}\left[\sup_{v\in[0,1]}\int_0^v|\sigma B(u)|\mathrm{d}u + \sup_{v\in[0,1]}\int_v^1|\sigma B(u)|\mathrm{d}u\right] \\
&= \frac{1}{T}I_T(B) \\
&= o_P(1),
\end{aligned}$$

用相同的方法可以得到，当 $T\to\infty$时

$$\sup_{s\in[\kappa,1]}\left|Z_T(s)-\sigma^2\int_0^s B^2(r)\mathrm{d}r\right| = o_P(1),$$

综上所述得到，当 $T\to\infty$时

$$\sup_{s\in[\kappa,1]}\left\|\begin{pmatrix}X_T(s)\\ Z_T(s)\end{pmatrix}-\begin{pmatrix}\zeta\sigma^2\int_0^s\left(\int_0^r B(u)\mathrm{d}u\right)^2 K(\zeta(r-s))\mathrm{d}r\\ \sigma^2\int_0^s B^2(r)\mathrm{d}r\end{pmatrix}\right\| = o_P(1),$$

其中$\|\cdot\|$表示 $\mathbf{R}^2$ 中的任一向量范数. 而上式表明了原始过程在乘积空间 $D([0,1];\mathbf{R})\times D([0,1];\mathbf{R})$(配备乘积 Skorohod 拓扑，见附录 B. 2)上的联合弱收敛

$$(X_T,Z_T)\Rightarrow(X,Z),\quad T\to\infty,$$

其中

$$X(s) = \zeta\sigma^2\int_0^s\left(\int_0^r B(u)\mathrm{d}u\right)^2 K(\zeta(r-s))\mathrm{d}r,$$

$$Z(s) = \sigma^2\int_0^s B^2(r)\mathrm{d}r,$$

377

$s\in[\kappa,1]$. 接下来，注意到：对任意 $s\in[\kappa,1]$，有

$$P(Z(s)>0)=1,$$

否则 $B(r)=0$(在$[\kappa,1]$中 λ-几乎处处)的概率为正，这是一个矛盾；在任意无限区间上布朗运动的方差是无限的. 因此，由连续映射定理(见评注 B. 2. 6)，有：当 $T\to\infty$时

$$U_T = \frac{X_T}{Z_T}\Rightarrow\mathcal{U}.$$

由上面的定理，可以推出规范停时法则

$$\frac{R_T}{T}=\inf\left\{\kappa\leqslant\frac{t}{T}\leqslant 1: U_T(s)<c\right\}$$

的中心极限定理．注意到：R_T/T 晚于 x 停止，当且仅当条件 $U_T(s)\geqslant c$ 对所有 $s\in[\kappa,\ x]$ 成立．这意味着

$$\frac{R_T}{T}>x \quad\Leftrightarrow\quad \sup_{s\in[\kappa,x]}U_T(s)\geqslant c,$$

对所有给定的 $x\in\mathbf{R}$ 成立．因此，对 R_T/T 的分布函数 $F_T(x)$，当 $T\to\infty$ 时有

$$\begin{aligned}F_T(x)&=P(R_T/T\leqslant x)\\&=P(\sup_{s\in[\kappa,x]}U_T(s)<c)\\&\to P(\sup_{s\in[\kappa,x]}\mathcal{U}_1(s)<c).\end{aligned}$$

前面已经提及到：在一个给定的显著水平下，为了建立大样本下检验和监测过程，上述结果使人们能够分别得到临界值和控制极限．

I(1)检测：为了检测从 $I(0)$ 到 $I(1)$ 的变点，考虑过程

$$\widetilde{U}_T(s)=\frac{1}{Ts_{Tm}^2(s)}\sum_{i=1}^{\lfloor Ts\rfloor}\left(\sum_{j=1}^{i}Y_j\right)^2K_h(i-\lfloor Ts\rfloor),\quad s\in[0,1],\tag{9.35}$$

其中

$$s_{Tm}^2(s)=\frac{1}{T}\sum_{i=1}^{\lfloor Ts\rfloor}Y_i^2+2\sum_{k=1}^{m}w(k,m)\frac{1}{T}\sum_{i=1}^{\lfloor Ts\rfloor}Y_iY_{i+k},\quad s\in[0,1]$$

378

是由(8.28)式定义的在 8.8 节中的定理 8.8.6 中讨论过的 Newey-West 估计量．该估计量排除了极限分布中可能出现的麻烦参数．对应的检测因子为

$$\widetilde{R}_T=\widetilde{R}_T(c)=\min(k\leqslant n\leqslant T:\widetilde{U}_T(n/T)<c).$$

该检测因子在原假设下的分布由下面定理给出．

定理 9.4.9 设 $\{Y_n\}$ 是一个严平稳的 α 混合 $I(0)$ 过程，满足 $EY_1^{4\nu}<\infty$，且 α 混合系数满足对某个 $\nu>1$，

$$\sum_{j=1}^{\infty}j^2\alpha(j)^{(\nu-1)/\nu}<\infty.\tag{9.36}$$

假设滞后截断法则满足 $m/T^{1/2}=o(1)$．那么，当 $T\to\infty$ 时在 $D[\kappa,\ 1]$ 中

$$\widetilde{U}_T(s)\Rightarrow\widetilde{\mathcal{U}}_2(s)=s^{-1}\zeta\int_0^s B(r)^2K(\zeta(r-s))\mathrm{d}r.$$

现在考虑近似随机游动时间序列的渐近性．考虑下列数学模型：在分析三角阵列 $\{Y_{Tt}:1\leqslant tT,\ T\in\mathbf{N}\}$ 的第 T 行时，已经得到了前 T 个观测值，且

$$Y_{T,t+1}=(1+a/T)Y_{Tt}+u_t,\quad 1\leqslant t\leqslant T, T\in\mathbf{N},\tag{9.37}$$

其中 $a\in\mathbf{R}$ 且 $\{u_t\}$ 是一个 $I(0)$ 过程．虽然对每个固定的 T，序列 Y_{T1}，…，Y_{TT} 是 $I(0)$ 平稳的，但是，当 $T\to\infty$ 时，AR 系数 $1+a/T$ 趋于 1(等于 1 是单位根的情形)．这个模型称为**局部一致模型**．

下面简单地讨论一下这个比较一般的模型的渐近性．由微积分理论知

$$\frac{1}{\sqrt{T}}Y_{T,\lfloor Ts\rfloor}=\frac{1}{\sqrt{T}}\sum_{t=1}^{\lfloor Ts\rfloor}(1+a/T)^{\lfloor Ts\rfloor-i}u_t$$
$$=\int_0^{\lfloor Ts\rfloor/T}e_T(r;s)\mathrm{d}S_T(r),$$

其中

$$S_T(r)=\frac{1}{\sqrt{T}}\sum_{t=1}^{\lfloor Ts\rfloor}u_t,$$

379

$$e_T(r;s)=(1+a/T)^{\lfloor Ts\rfloor-\lfloor Tr\rfloor},$$

r，$s\in[0，1]$. 注意到：当 $0\leqslant r\leqslant s\leqslant 1$ 时，$e_T(r;\ s)$一致收敛于 $e(r;\ s)=e^{a(s-r)}$，并且是一致有界变差函数，从而得到：当 $T\to\infty$时

$$\sup_{s\in[0,1]}\left|\int_0^{\lfloor Ts\rfloor/T}e_T(r;s)\mathrm{d}S_T(r)-\int_0^s e(r;s)\mathrm{d}B(r)\right|\to 0(依概率).$$

这就得到了下列结论.

命题 9.4.10(收敛于 Ornstein-Uhlenbeck 过程) 设$\{Y_{Tt}：1\leqslant t\leqslant T，T\geqslant 1\}$是一个满足局部-一致模型(9.37)的矩阵序列，则当 $T\to\infty$ 时

$$\frac{1}{\sqrt{T}}Y_{T,\lfloor Ts\rfloor}=\int_0^{\lfloor Ts\rfloor/T}e_T(r;s)\mathrm{d}S_T(r)\Rightarrow\mathcal{O}(r;s),$$

其中

$$\mathcal{O}(r;s)=\int_0^s \mathrm{e}^{a(s-r)}\mathrm{d}B(r)$$

是一个 Ornstein-Uhlenbeck 过程，B 是标准布朗运动.

在定理 9.4.8 的证明中，将条件改为一般性假设：标准过程 $T^{-1/2}Y_{T,\lfloor Ts\rfloor}$，$s\in[\kappa，1]$弱收敛于一个几乎必然连续过程，证明过程仍可进行下去，通过简单的修正可以得到数据生成过程的渐近性质，从而得到更一般的极限结果. 因此，可以得到下列过程的**局部一致渐近**性质

$$U_T(s)=\frac{\lfloor Ts\rfloor^{-3}\sum_{i=1}^{\lfloor Ts\rfloor}\left(\sum_{j=1}^{i}Y_{Tj}\right)^2K_h(i-\lfloor Ts\rfloor)}{\lfloor Ts\rfloor^{-2}\sum_{j=1}^{\lfloor Ts\rfloor}Y_{Tj}^2},s\in[1/T,1]. \tag{9.38}$$

定理 9.4.11 设$\{Y_{Tt}：1\leqslant t\leqslant T，T\geqslant 1\}$是一个满足局部一致模型(9.37)的矩阵序列. 则当 $T\to\infty$时，在 $D[\kappa，1]$中，

$$U_T(s)\Rightarrow\mathcal{U}_{\mathcal{O}}(s)=\frac{\zeta s^{-1}\int_0^s\left(\int_0^r\mathcal{O}(t;a)\mathrm{d}t\right)^2K(\zeta(s-r))\mathrm{d}r,}{\int_0^s\mathcal{O}(r;a)^2\mathrm{d}r}.$$

注意到，过程 $U_t(s)$在比例改变时是不变的. 因此可以考虑强相依过程，它们有不同的收敛速度. 设

$$Y_t = \sum_{i=1}^{t} X_i, \quad t = 1,2,\cdots \tag{9.39}$$

有长记忆增量 X_t，比如，它是一个阶数为 $0<d<1/2$ 的分数单整噪声，其 Hurst 指数为 $H=1/2+d$，即

$$(1-L)^d X_t - \mathrm{WN}(0,\sigma^2), \quad \sigma^2 \in (0,\infty),$$

380

见定义 3.8.3. 下面，设$\{X_t\}$满足极限带分数布朗运动的不变原理：当 $T\to\infty$时，

$$\frac{1}{T^H}\sum_{t=1}^{\lfloor Ts \rfloor} X_t \Rightarrow c_H B^H(s), \tag{9.40}$$

对某个常数 c_H，见定理 B.7.5 及参考文献. 从而，标准过程乘以 T^{-H}满足

$$\frac{1}{T^H} Y_{\lfloor Ts \rfloor} \Rightarrow c_H B^H, \quad T \to \infty,$$

这就得到基于时间序列(9.39)的过程 $U_T(s)$满足：当 $T\to\infty$时，

$$U_T(s) \Rightarrow \mathcal{U}_{B^H}(s) = \frac{\zeta s^{-1}\int_0^s \left(\int_0^r B^H(t)\mathrm{d}t\right)^2 K(\zeta(s-r))\mathrm{d}r,}{\int_0^s B^H(r)^2 \mathrm{d}r}.$$

9.5 评注与延伸阅读

Sklar 定理可以在 Sklar(1959)中查到，也可以参考 Rüschenorf(2009). 风险管理中的相关性和相依性讨论及 copula 函数的性质可以参考 Eebrechts(2002)等. 金融危机的简要概述主要选自 Baily 和 Johnson(2008). 应用高斯 copula 函数对 CDO 定价选自 Li(2000). 非参数 copula 估计量的渐近性质参考了 Fermanian(2004)等. 首先对局部多项式回归进行系统研究的是 Stone(1977)，要深入学习这一理论及应用方面的有关知识，可参考 Fan 和 Gijbels(1996)，Fan 和 Yao(2003)；Li 和 Racine(2007)从计量经济学的角度介绍了这一理论及应用. 有关极小极大估计的内容可以在 Tsybabov(2009)中找到. Masry 和 Fan(1997)建立了混合过程的渐近性质，读者也可以参考 Fan 和 Yao(2003)以及 Li 和 Racine(2007). Bosq(1998)系统地介绍了相依随机过程的非参数统计量的有关知识. 要学习非参数计量经济学的方法，特别是核平滑方法，我们推荐专著 Li 和 Racine(2007). 本章的内容参考了上述著作，并给出了渐近正态性的全新形式，得出了对任意满足平滑中心极限定理的平稳遍历时间序列，渐近正态性结论都成立.

Neyman-Pearson 型结论(定理 9.3.2)及应用可参考 Vexler 和 Gurevich(2011). 定理 9.3.7 及相关的结果可以在 Steland(2004)中找到. 最优核理论请参考 Steland(2005). 关于变点问题的系统论述，我们推荐专著 Csorgo 和 Horvath(1997)，本章主要集中于检测平稳性与非平稳性的序贯方法. 本章讨论的基于连续 KPSS 过程的 I(0)平稳性和 I(1)非平稳性的序贯方法的渐近性结论，读者可参考 Steland(2007b)，或 Davies 和 Kramer(2003). Steland(2007b)研究了 Dickey-Fuller 单位根检验统计量，Steland(2008)则将其推广到有多项

381

式趋势的情形. 在线性模型中检测变点的方法，我们推荐 Ploberger 和 Kramer(1992) 和 Huskova(2007)等. 使(9.40)式成立的条件，可以参考 Wu 和 Shao(2006).

参考文献

Baily, Martin Neil L.R.E and Johnson M.S. (2008) The origins of the financial crisis. The Brookings Institution, pp. 1–45.

Bosq D. (1998) *Nonparametric Statistics for Stochastic Processes: Estimation and Prediction*. vol. 110 of *Lecture Notes in Statistics* 2nd edn. Springer-Verlag, New York.

Csörgő M. and Horváth L. (1997) *Limit Theorems in Change-point Analysis*. Wiley Series in Probability and Statistics. John Wiley & Sons Ltd., Chichester. With a foreword by David Kendall.

Davies P.L. and Krämer W. (2003) The Dickey-Fuller test for exponential random walks. *Econometric Theory* **19**(5), 865–883.

Embrechts P., McNeil A.J. and Straumann D. (2002) Correlation and dependence in risk management: properties and pitfalls *Risk Management: Value at Risk and Beyond (Cambridge, 1998)*. Cambridge Univ. Press Cambridge pp. 176–223.

Fan J. and Gijbels I. (1996) *Local Polynomial Modelling and its Applications*. vol. 66 of *Monographs on Statistics and Applied Probability*. Chapman & Hall, London.

Fan J. and Yao Q. (2003) *Nonlinear Time Series: Nonparametric and Parametric Methods*. Springer Series in Statistics. Springer-Verlag, New York.

Fermanian J.D., Radulović D. and Wegkamp M. (2004) Weak convergence of empirical copula processes. *Bernoulli* **10**(5), 847–860.

Hušková M., Prášková Z. and Steinebach J. (2007) On the detection of changes in autoregressive time series. I. Asymptotics. *J. Statist. Plann. Inference* **137**(4), 1243–1259.

Li D.X. (2000) On default correlation: A copula function approach. *The Journal of Fixed Income* **9**, 43–54.

Li Q. and Racine J.S. (2007) *Nonparametric Econometrics*. Princeton University Press, Princeton, NJ. Theory and practice.

Masry E. and Fan J. (1997) Local polynomial estimation of regression functions for mixing processes. *Scand. J. Statist.* **24**(2), 165–179.

Ploberger W. and Krämer W. (1992) The CUSUM test with OLS residuals. *Econometrica* **60**(2), 271–285.

Rüschendorf L. (2009) On the distributional transform, Sklar's theorem, and the empirical copula process. *J. Statist. Plann. Inference* **139**(11), 3921–3927.

Sklar M. (1959) Fonctions de répartition à n dimensions et leurs marges. *Publ. Inst. Statist. Univ. Paris* **8**, 229–231.

Steland A. (2004) Sequential control of time series by functionals of kernal-weighted empirical processes under local alternatives. *Metrika* **60**(3), 229–249.

Steland A. (2005) Optimal sequential kernel detection for dependent processes. *J. Statist. Plann. Inference* **132**(1-2), 131–147.

Steland A. (2007a) Monitoring procedures to detect unit roots and stationarity. *Econometric Theory* **23**(6), 1108–1135.

Steland A. (2007b) Weighted Dickey-Fuller processes for detecting stationarity. *J. Statist. Plann. Inference* **137**(12), 4011–4030.

Steland A. (2008) Sequentially updated residuals and detection of stationary errors in polynomial regression models. *Sequential Anal.* **27**(3), 304–329.

Stone C.J. (1977) Consistent nonparametric regression. *Ann. Statist.* **5**(4), 595–645. With discussion and a reply by the author.

Tsybakov A.B. (2009) *Introduction to Nonparametric Estimation*. Springer Series in Statistics. Springer, New York. Revised and extended from the 2004 French original, Translated by Vladimir Zaiats.

Vexler A. and Gurevich G. (2011) A note on optimality of hypothesis testing. *Mathematics in Engineering, Science and Aerospace* **2**(3), 243–250.

Wu W.B. and Shao X. (2006) Invariance principles for fractionally integrated nonlinear processes *Recent Developments in Nonparametric Inference and Probability* vol. 50 of *IMS Lecture Notes Monogr. Ser.* Inst. Math. Statist. Beachwood, OH pp. 20–30.

382 ~ 383

附　录　A

A.1　(随机)Landau 记号

对一个确定性的序列$\{a_n: n\in\mathbf{N}\}$，如满足：当 n 充分大时，$|a_n|\leqslant C$(即存在某个 $n_0\in\mathbf{N}$，使得当$n\geqslant n_0$ 时，有$|a_n|\leqslant C$)，则记为 $a_n=O(1)$．如果当 $n\to\infty$时，有 $a_n\to 0$，则记 $a_n=o(1)$．两种记号都可推广到向量序列上去，更一般地，可推广到赋范线性空间上去．对实值序列$\{b_n\}$，还可以使用下列记号：如果存在某个常数 $C<\infty$，使得当 $n\to\infty$时，有$|a_n|\leqslant C|b_n|$，即$\lim\limits_{n\to\infty}\left|\dfrac{a_n}{b_n}\right|<\infty$，则记 $a_n=O(b_n)$；如果当$n\to\infty$时，有 $a_n/b_n\to 0$，即 $\limsup\limits_{n\to\infty}\left|\dfrac{a_n}{b_n}\right|=0$，则记 $a_n=o(b_n)$．

设$\{X_n: n\in\mathbf{N}\}$是一个随机变量序列，如果对任意 $\varepsilon>0$，存在常数 $M>0$ 和 $n_0\in\mathbf{N}$(它通常与 ε 有关)，使得当 $n\geqslant n_0$ 时，有

$$P(|X_n|>M)\leqslant\varepsilon,$$

则称$\{X_n\}$是依概率有界的，记为 $X_n=O_p(1)$．如果扩大 M，则对有限项 $n=1,\cdots,n_0$，$P(|X_n|>M)\leqslant\varepsilon$ 也是成立的，所以上面的定义等价于

$$\sup_{n\in\mathbf{N}}P(|X_n|>M)\leqslant\varepsilon.$$

依概率有界的序列$\{X_n\}$也称为一致紧的．$X_n=O_P(1)$的一个充分条件为 X_n 依分布收敛于某个非减的随机变量或 $\mathbf{R}$ 中某个点 x_0．如果当 $n\to\infty$时，有 $X_n\xrightarrow{P}0$，则记为

$$X_n=o_P(1).$$

上述两种记号均能推广到随机向量的情形，也适用于随机赋范空间．如果$\{Y_n\}$是一个实值随机变量序列，且满足

$$X_n=Y_nR_n,\text{其中 } R_n=O_P(1),$$

则记 $X_n=O_P(Y_n)$．如果 $Y_n\neq 0$，且 $X_n/Y_n=O_p(1)$，则有 $X_n=O_P(Y_n)$．如果

$$\frac{X_n}{Y_n}=o_P(1),\text{即当 } n\to\infty \text{ 时,有}\left|\frac{X_n}{Y_n}\right|\xrightarrow{P}0,$$

则记

$$X_n=o_P(Y_n).$$

记号 $O_P(1)$与 $o_P(1)$所具有的性质是非常重要的，这种项在不同的地方其具体表达式是不同的．使用这种记号通常能使计算与证明变得简单．

随机阶数可以通过估计矩而获得：

(ⅰ) 如果 $E(\|X_n\|)=O(a_n)$，那么 $X_n=O_P(a_n)$；

(ⅱ) 如果 $E(\|X_n\|^2)=O(a_n)$，那么 $X_n=O_P(a_n^{1/2})$．

下面是 O_P 与 o_P 的计算法则：

(ⅰ) $o_P(1)+o_P(1)=o_P(1)$；

(ⅱ) $O_P(1)+o_P(1)=O_P(1)$；

(ⅲ) $o_P(1)O_P(1)=o_P(1)$；

(ⅳ) $o_P(R_n)=R_no_P(1)$；

(ⅴ) $O_P(R_n)=R_nO_P(1)$;

(ⅵ) $o_P(O_P(1))=o_P(1)$.

下列重要引理表明：在 $o(\cdot)$、$O(\cdot)$这种余项中，比如出现在 Taylor 展式中的余项，可以插入 $O_P(1)$与 $o_P(1)$这种项.

引理 A.1.1 设 $f: D\to\mathbf{R}$ 是一个定义在 $D\subset\mathbf{R}^k$ 上的函数，其中 $k\in\mathbf{N}$，且$\{X_n\}$是一个取值在 D 中的随机变量序列，满足 $X_n=o_P(1)$，则

(ⅰ) 如果当 $x\to 0$ 时，$f(x)=o(\|x\|^r)$,，那么 $f(X_n)=o_P(\|X_n\|^r)$.

(ⅱ) 如果当 $x\to 0$ 时，$f(x)=O(\|x\|^r)$，那么 $f(X_n)=O_P(\|X_n\|^r)$.

证明见 Vander Vaart(1998).

386

A.2 Bochner 引理

引理 A.2.1 设$\{g_n: n\geqslant 1\}$，g_n: $\mathbf{R}\to\mathbf{R}$ 是一致可积的函数列，即

$$\sup_{n\in\mathbf{N}}\int |g_n(y)|\mathrm{d}y\leqslant C<\infty,$$

且在任意紧集上一致收敛于函数 g，即对任一紧集 $I\subset\mathbf{R}$，有

$$\lim_{n\to\infty}\sup_{x\in I}|g_n(x)-g(x)|=0,$$

进一步设 K: $\mathbf{R}\to\mathbf{R}$ 是一个满足下列条件的函数：

(ⅰ) $\int K(z)\mathrm{d}z<\infty$,

(ⅱ) 当 $|z|\to\infty$时，$|z||K(z)|\to 0$.

序列 h_n 满足 $h_n\downarrow 0$，那么，光滑算子

$$\widetilde{g}_n(x)=K_{h_n}\star g_n(x)=\int h_n^{-1}K([x-y]/h_n)g_n(y)\mathrm{d}y$$

满足：对任意紧集 $I\subset\mathbf{R}$，均有

$$\lim_{n\to\infty}\sup_{x\in I}\left|\widetilde{g}_n(x)-g(x)\int K(z)\mathrm{d}z\right|=0$$

成立. ■

A.3 条件期望

设 $X:(\Omega, \mathcal{F}, P)\to(\mathbf{R}, \mathcal{B})$是一个随机变量，首先回忆一下期望的定义：如果 $\min\{E(X^+), E(X^-)\}<\infty$，则称$\int X\mathrm{d}P$ 存在或有定义，记 $E(X)=E(X^+)-E(X^-)$，如果 $E(|X|)<\infty$，或等价地 $E(X^+)<\infty$及 $E(X^-)<\infty$，则称 $E(X)$是有限的. 由$|x|=x^+ + x^-$知 $E(X)<\infty$($E(X)$是有限的)当且仅当$E(|X|)<\infty$.

现设$\mathcal{A}\subset\mathcal{F}$ 是一个子σ代数，如果 $X\geqslant 0$，则映射 $A\mapsto\int_A X\mathrm{d}P$，$A\in\mathcal{F}$定义了$\mathcal{A}$上的一个测度 υ，如果 $EX<\infty$，则这个测度是有限的. Radon-Nikodym 定理保证了在任何情况下都存在一个 $\mathcal{A}$ 可测的密度 f，使得$\upsilon(A)=\int_A f\mathrm{d}P$，$A\in\mathcal{A}$，称 f 为 X 在条件 $\mathcal{A}$ 下的条件期望，记为 $E(X|\mathcal{A})$. 如果 $E(X)<\infty$，那么 $E(X|\mathcal{A})<\infty$. 假定 X 在 $\mathbf{R}$ 上取值，且满足 $E(|X|)<\infty$，将 X 分解为 $X=X^+-X^-$，其中 X^+，$X^-\geqslant 0$，用 f^+，f^-表示对应的密度，那么，条件期望 $E(X|\mathcal{A})$是$\mathcal{A}$可测的函数 f^+-f^-，且满足：$\int_A E(X|\mathcal{A})\mathrm{d}P=$

$\int_A X\mathrm{d}P$. 注意，当$|X|$不可积时，条件期望也可能有定义，例如，见 Shiryaev(1999). 如按上述方法已经定义了$E(X^+|\mathcal{A})$与$E(X^-|\mathcal{A})$，设

387

$$\min\{E(X^+|\mathcal{A})(\omega), E(X^-|\mathcal{A})(\omega)\} < \infty \tag{A.1}$$

对(几乎)所有$\omega\in\Omega$成立，那么可以定义X的广义条件期望为$E(X|\mathcal{A})=E(X^+|\mathcal{A})-E(X^-|\mathcal{A})$，其中在零测集$\{E(X^+|\mathcal{A})=E(X^-|\mathcal{A})=\infty\}$上. $E(X|\mathcal{A})$可以定义为任意值，例如0. 如果$E(|X||\mathcal{A})(\omega)<\infty$对(几乎)所有$\omega\in\Omega$成立，则公式(A.1)成立，且对(几乎)所有$\omega\in\Omega$，$E(X|\mathcal{A})(\omega)$是有限的，称广义条件期望 a.s. 存且有限.

定理 A.3.1 条件期望的性质如下.

设$\mathcal{A}\subset\mathcal{F}$是一个子$\sigma$代数，只要下列所有随机变量的期望存在，则下列结论均成立.

(i) 如果$\mathcal{A}=\{\varnothing, \Omega\}$，那么$E(X|\mathcal{A})=E(X)$(a.s.).

(ii) 如果$X=c$ (a.s.)，那么$E(X|\mathcal{A})=c$ (a.s.).

(iii) 如果$X\leqslant Y$ (a.s.)，那么$E(X|\mathcal{A})\leqslant E(Y|\mathcal{A})$(a.s.).

(iv) $|E(X|\mathcal{A})|\leqslant E(|X||\mathcal{A})$(a.s.).

(v) 设a，b是常数，且$aE(X)+bE(Y)$存在，那么$E(aX+bY|\mathcal{A})=aE(X|\mathcal{A})+bE(Y|\mathcal{A})$.

(vi) $E(X)=E(E(X|\mathcal{A}))$.

(vii) 如果$\mathcal{A}_1\subset\mathcal{A}_2$，那么$E(E(X|\mathcal{A}_2)|\mathcal{A}_1)=E(X|\mathcal{A}_1)$(a.s.)，

如果$\mathcal{A}_2\subset\mathcal{A}_1$，那么$E(E(X|\mathcal{A}_2)|\mathcal{A}_1)=E(X|\mathcal{A}_2)$(a.s.).

(viii) 如果X独立于$\mathcal{A}$，且$E(X)$存在，那么$E(X|\mathcal{A})=E(X)$(a.s.).

(ix) 设X和Y是两个随机变量，且X是$\mathcal{A}$可测的，$E|X|<\infty$，$E|XY|<\infty$，那么$E(XY|\mathcal{A})=XE(Y|\mathcal{A})$(a.s.).

(x) 如果在集合$A\in\mathcal{A}$上$X=Y$，那么$\mathbf{1}_A E(Y|\mathcal{A})+\mathbf{1}_{A^c}E(X|\mathcal{A})$是$E(X|\mathcal{A})$的一个修正，即对$\omega\in\mathcal{A}$，可用$E(Y|\mathcal{A})(\omega)$代替$E(X|\mathcal{A})(\omega)$.

(xi) 如果φ: $\mathbf{R}\to\mathbf{R}$是凸的，且$E|\phi(X)|<\infty$，那么$\varphi(E(X|\mathcal{A}))\leqslant E(\varphi(X)|\mathcal{A})$. ■

A.4 不等式

下列定理列举了本书常用的一些经典不等式.

定理 A.4.1 设X是取值于集合χ的随机变量.

(i) Markov 不等式：设g: $\chi\to\mathbf{R}$是可测函数，且满足$Eg(X)<\infty$，那么对任意常数$c>0$，有$p(X>c)\leqslant E(g(X))/g(c)$.

388

(ii) 切比雪夫不等式：如$E(X^2)<\infty$，那么对任意$\varepsilon>0$，有$P_o(|X-E(X)|>\varepsilon)\leqslant \mathrm{Var}(X)/\varepsilon^2$.

(iii) Hölder 不等式：设$X\in L_p$，$Y\in L_q$，其中p，$q\in[0, \infty]$，$\frac{1}{p}+\frac{1}{q}=1$，那么$E|XY|\leqslant [E(X^p)]^{1/p}[E(Y^q)]^{1/q}$，即$\|XY\|_{L_1}\leqslant\|X\|_{L_p}\|Y\|_{L_q}$.

(iv) 广义 Hölder 不等式：设X_1，…，X_n是一列随机变量，满足$X_i\in L_{p_i}(i=1, \cdots, n)$，其中$p_1$，…，$p_n\in(0, \infty)$，且对$r\in(0, \infty)$有$\frac{1}{p_1}+\cdots+\frac{1}{p_n}=\frac{1}{r}$，那么

$$\left\|\prod_{i=1}^n X_i\right\|_r \leqslant \prod_{i=1}^n \|X_i\|_{p_i}.$$

（v）C_r 不等式：设 X 和 Y 是两个随机变量，满足 $E(X^r)<\infty$，$E(Y^r)<\infty$，那么，

$$E|X+Y|^r \leqslant C_r[E|X|^r + |Y|^r],$$

其中，如果 $0<r\leqslant 1$，则 $C_r=1$，如果 $r>1$，则 $C_r=2^{r-1}$.

A.5　Random 序列

下列定理确保了一类线性过程的存在性.

定理 A.5.1　设$\{X_n: n\in\mathbf{N}\}$是一列 L_1 有界的随机变量，即 $\sup\limits_{n\in\mathbf{N}} E|X_n|<\infty$. 设$\{a_n: n\in\mathbf{N}\}\subset\mathbf{C}$ 是可求和序列，即 $\sum\limits_{n=1}^{\infty}|a_n|<\infty$，那么，随机变量序列

$$Y_n = \sum_{i=0}^{\infty} a_i X_{n-i}, \quad n\in\mathbf{N}, \tag{A.2}$$

几乎必然存在. 如果$\{X_n\}$是 L_2 有界的，即$\sup\limits_{n\in\mathbf{N}} EX_n^2<\infty$，那么，(A.2)在 L_2 意义下成立，且有

$$E\Big(\sum_{i=0}^{\infty} a_i X_{n-i}\Big) = \sum_{i=0}^{\infty} a_i E(X_{n-i}), \tag{A.3}$$

$$E\Big(\sum_{i=0}^{\infty} a_i X_{n-i}\Big) = \lim_{N\to\infty} E\Big(\sum_{i=0}^{N} a_i X_{n-i}\Big). \tag{A.4}$$

显然，将本定理中的 $\mathbf{N}$ 换成 $\mathbf{Z}$ 后，结论仍成立.

A.6　离散时间的局部鞅

称过程$\{X_t\}$为局部鞅，如果存在一列满足当 $n\to\infty$时，$T_n\uparrow\infty$的停时$\{T_n\}$，使得停时化过程 $X^{T_n}=\{X_{\min(t,T_n)}: t\geqslant 0\}(n\geqslant 1)$是一个鞅，即 $E|X^{T_n}|<\infty$，且对所有 $s\leqslant t$ 和 $n\geqslant 1$，均有 $E(X_t^{T_n}|\mathcal{F}_s)=X_s^{T_n}$ 成立.

389

设$\{X_t: t\geqslant 0\}$是定义在过滤概率空间$(\Omega, \mathcal{F}, \{\mathcal{F}_t\}, P)$上的右连左极过程，如果它有下列分解，

$$X_t = X_0 + A_t + M_t, \quad t\geqslant 0,$$

其中，$\{A_t\}$为有界变差过程，$\{M_t\}$为局部鞅，则称$\{X_t\}$为一个**半鞅**.

对离散时间的情形，有下列性质(见 Jacod 和 shiryaev(2003)命题 1.4.6).

定理 A.6.1　下列条件是等价的：

（i）$\{X_t: t\in\mathbf{N}\}$是一个局部鞅；

（ii）$\{X_t: t\in\mathbf{N}\}$是一个鞅变换，即存在一个鞅$\{M_t: t\in\mathbf{N}\}$和一个可预报过程$\{H_t: t\in\mathbf{N}\}$，使得对任意 $n\geqslant 1$，有 $X_n=X_0+\sum\limits_{i=0}^{n} H_{i-1}(M_i-M_{i-1})$；

（iii）$\{X_t: t\in\mathbf{N}\}$是一个广义鞅，即对任意 $t\geqslant 1$，有 $E(|X_t|\mathcal{F}_{t-1})<\infty$，且 $E(X_t|\mathcal{F}_{t-1})=X_{t-1}$，a.s..　■

390

附录B　弱收敛与中心极限定理

B.1　依分布收敛

设

$$X_n:(\Omega_n,\mathcal{F}_n,P_n)\to(\mathbf{R},\mathcal{B}),\quad n\in\mathbf{N}$$

是一列随机变量，其分布函数分别为 $F_n(x)=P(X_n\leqslant x)$，$x\in\mathbf{R}$，$n\in\mathbf{N}$. 从下面的讨论可看出，对不同的 $n\in\mathbf{N}$，X_n 还可以分别属于不同的概率空间$(\Omega_n,\mathcal{F}_n,P_n)$，又设 F 为一个新的分布函数，$X\sim F$，如果对 F 在 $\mathbf{R}$ 上的所有连续点 x，即 $x\in C_F=\{y\in\mathbf{R}: F(y-)=F(y)\}$，均有

$$F(x)=\lim_{n\to\infty}F_n(x),$$

则称$\{F_n\}$**依分布收敛**于 F. 显然依分布收敛的定义可以直接推广到 $\mathbf{R}^d$ 上的分布函数上去.

称随机变量或随机向量$\{X_n\}$**依分布收敛**于 X，如果$\{X_n\}$对应的分布函数列$\{F_n\}$依分布收敛于 X 的分布函数 F，记为 $X_n\xrightarrow{d}X$，$n\to\infty$. 由定义不难看出，对所有 $x\in C_F$，均有

$$P(X_n\leqslant x)\to P(X\leqslant x),\qquad n\to\infty.$$

依分布收敛有下列等价定义：设 $\varphi_n(t)=E(\mathrm{e}^{itX_n})$ 与 $\varphi(t)$分别为 X_n 与 X 的特征函数(ch. f.)，$t\in\mathbf{R}$，则当 $n\to\infty$时$X_n\xrightarrow{d}X$ 当且仅当对所有 $t\in\mathbf{R}$，均有

$$\varphi_n(t)\to\varphi(t),n\to\infty.$$

由特征函数的收敛性马上得到下列定理，它是建立高维中心极限定理的重要工具.

定理 B.1.1(CRAMÉR-WOLD 工具)　维数为 $d\in\mathbf{N}$ 的随机向量列$\{X, X_n\}$依分布收敛，即当 $n\to\infty$ 时，$X_n\xrightarrow{d}X$，当且仅当对所有 $\lambda\in\mathbf{R}^d$，均有

$$\lambda' X_n\xrightarrow{d}\lambda' X,$$

当 $n\to\infty$时成立. ■

特别地，由 Cramer-Wold 工具知：对 $\mu\in\mathbf{R}^d$，$\Sigma\in\mathbf{R}^{d\times d}$，

$$X_n\xrightarrow{d}N(\mu,\Sigma),\quad n\to\infty,$$

当且仅当对每一个固定的向量 $\lambda=(\lambda_1,\cdots,\lambda_d)'\in\mathbf{R}^d$，一维随机变量序列 $Y_n=Y_n(\lambda)=\sum_{j=1}^{d}\lambda_j X_{nj}$ 满足形如定理 B.7.2 或定理 B.7.3 的一维中心极限定理. 下列定理确保了依分布收敛在连续变换下的不变性.

定理 B.1.2(连续映射)　设$\{X, X_n\}$是每一个分量均以概率 1 取值于 $\chi\subset\mathbf{R}^d$ 的 $d\in\mathbf{N}$ 维随机向量列，且

$$X_n\xrightarrow{d}X,\quad n\to\infty,$$

如果 ϕ: $\chi\to\mathbf{R}^m$，$m\in\mathbf{N}$ 是连续函数，则有

$$\varphi(X_n)\xrightarrow{d}\varphi(X),\quad n\to\infty$$

成立. ■

B.2　弱收敛

概率测度的弱收敛是将随机向量序列依分布收敛的概念推广到随机过程上去，或更一般地，推广到

取值于度量空间的随机元上去. 设(S, d)是一个度量为d的度量空间. 回顾一下，如果下列条件满足，称$d: S\times S\to \mathbf{R}$为一个**度量**.

（i）对所有$x, y\in S$，均有$d(x, y)\geq 0$.

（ii）对所有$x, y\in S$，均有$d(x, y)=d(y, x)$.

（iii）对所有$x, y\in S$，$d(x, y)=0$当且仅当$x=y$.

（iv）对所有$x, y, z\in S$，均有$d(x, z)\leq d(x, y)+d(y, z)$.

392

特别地，如果$(S, \|\cdot\|)$是一个赋范空间，则

$$d(x,y)=\|x-y\|,\quad x,y\in S,$$

是定义在S上的一个度量.

在S上配备由度量d确定的拓扑中的开集生成的Borel-σ代数$\mathcal{S}$. 如果该空间关于度量是完备的，即S中的任何Cauchy列在其内是收敛的，则可以使很多结论简化. 在数理金融中，可以通过选择适当的度量使基本利率空间是完备的.

下面是几个重要的度量空间，包括数列空间、连续函数空间、第一类不连续函数空间，在建模中，它们常被用来作为过程的路径.

（i）**时间序列**：设S为所有实数序列的集合，即

$$S=\mathbf{R}^{\infty}=\{\{x_n\}:X_n\in\mathbf{R},n\in\mathbf{R}\},$$

在其上配备度量

$$d(\{x_n\},\{y_n\})=\sum_{i=1}^{\infty}2^{-k}d_0(x_k,y_k),$$

其中，

$$d_0(x,y)=\frac{|x-y|}{1+|x-y|},\quad x,y\in\mathbf{R}.$$

则$S=\mathbf{R}^{\infty}$是一个可分的空间，因为$\forall x\in S$，存在$\{x_n\}\subset S$，对每个n，x_n只有有限项不为0，使得x_n依度量d收敛于x. 显然，可测映射$X: (\Omega, \mathcal{F}, P)\to(S, \mathcal{S})$是一维时间序列，因为如果用$\pi_n$表示$x=\{x_n\}$在第$n$个元素上的投影，即$\pi_n(x)=x_n$，那么$X=\{X_n\}$，其中$X_n=\pi_n\circ X$，即$X$是实值随机变量$X_n$的集合.

（ii）**有连续路径的随机过程**：设$S=C([0, 1];\mathbf{R})$为函数$f: [0, 1]\to\mathbf{R}$的集合，其中f关于通常的拓扑连续. 在其上配备由下列范数导出的一致度量

$$\|f\|=\|f\|_{\infty}=\sup_{t\in[0,1]}|f(t)|,\quad f\in C([0,1];\mathbf{R}).$$

（iii）**右连左极过程**：设$S=D([0, 1];\mathbf{R})$为右连左极函数$f: [0, 1]\to\mathbf{R}$的集合. 取值于S中的右连左极随机过程X的定义见定义5.1.2. 因此S的元素可能带跳，比较适合的度量是Skorohod度量

$$d'(f,g)=\inf_{\lambda\in\Lambda}\max\{\|f-g\circ\lambda\|_{\infty},\|\lambda-\mathrm{id}\|_{\infty}\},$$

393

其中，Λ表示所有连续且严格递增的映射λ: $[0, 1]\to[0, 1]$的集合. id表示$[0, 1]$上的恒同映射. 一个等价的度量为

$$d(f,g)=\inf_{\lambda\in\Lambda}\max\{\|f-g\circ\lambda\|_{\infty},\|\lambda\|^{\circ}\},$$

其中，对定义在$[0, 1]$上且满足$\lambda(0)=0$，$\lambda(1)=1$的非减函数，记

$$\|\lambda\|^{\circ}=\sup_{s<t}\left|\log\frac{\lambda(t)-\lambda(s)}{t-s}\right|.$$

如果$\|\lambda\|^{\circ}<\infty$，则$\lambda$的弦的斜率有位于0到$\infty$之间的界. 所以在这种情况下，$\lambda$是连续且严格递增的.

一个序列关于度量 d 收敛当且仅当它关于度量 d' 收敛，且在度量 d 下是一个完备的度量空间(见后面的讨论).

(ⅳ) **定义在[0, 1]上，取值于 $\mathbf{R}^d$ 的右连左极函数**：$d\geqslant 2$ 时的讨论与 $d=1$ 的讨论是类似的，将前面的绝对值理解为 $\mathbf{R}^d$ 向量的范数，则前面的记号仍适用. 相应的空间记为

$$D([0,1];\mathbf{R}^d)=\{f:[0,1]\to\mathbf{R}^d \mid f \text{ 是右连左极的}\}.$$

(ⅴ) **定义在[0, ∞)上的右连左极函数**：在处理带时间坐标 $t\in[0,\infty)$ 的右连左极过程时，下列 Skorohod 度量是适合的：对 f，$g\in D([0,\infty);\mathbf{R})$，令

$$d(f,g)=\sum_{t=1}^{\infty}\frac{1}{2^t}\min\{1,d_t(f,g)\},$$

其中，$d_t(f,g)$按下列方式定义：$d_t(f,g)$是满足下列条件的 $\delta>0$ 的下确界，存在两个网格$\{t_i: i=1,\cdots,k\}$和$\{s_i: i=1,\cdots,k\}$，或满足 $t_0=s_0=0$，s_k，$t_k\geqslant t$ 的有序点，使得对 $i=1,\cdots,k$ 均有 $|t_i-s_i|\leqslant\delta$，且对 $i=0,\cdots,k-1$，均有

$$|f(t')-g(s')|\leqslant\delta,\quad \text{如果 } t'\in[t_i,t_{i+1}],s'\in[s_i,s_{i+1}],$$

见 Pollard(1984). 可以证明

$$d(f_n,f)\to 0,\quad n\to\infty,$$

当且仅当存在由$[0,\infty)$映到自己的连续、递增的映射$\{\lambda_n\}$，使得当 $n\to\infty$时，在紧集上一致地有

$$\lambda_n(t)-t\to 0,\quad \text{且 } f(\lambda_n(t))-f_n(t)\to 0.$$

对随机变量序列，依分布收敛是指

394

$$F_n(x)=\int\mathbf{1}_{(-\infty,x]}(s)\mathrm{d}F_n(s)\to\int\mathbf{1}_{(-\infty,x]}(s)\mathrm{d}F(s)=F(x),\quad n\to\infty,$$

对所有 $x\in C_F$ 成立. 用连续函数逼近示性函数，可以证明上式等价于

$$\int\varphi(s)\mathrm{d}F_n(s)\to\int\varphi(s)\mathrm{d}F(s),\quad n\to\infty, \tag{B.1}$$

对所有实值、连续、有界函数 φ，即 $\varphi\in C_b(\mathbf{R},\mathbf{R})$成立. (B.1)也可以理解为概率测度序列$\{\mathrm{d}F_n,\mathrm{d}F\}$在实际路径上的一个条件. 这个定义可以容易地推广到度量空间上的概率测度上去：称度量空间(S,d)上的概率测度$\{P_n,P\}$**弱收敛**，如果

$$\int\varphi(x)\mathrm{d}P_n(x)\to\int\varphi(x)\mathrm{d}P(x),\quad n\to\infty,$$

对所有 $\varphi\in C_b(S,\mathbf{R})$成立. 基于这个定义的许多基本计算法则，统计量在非标准条件下的极限分布，都可由连续映射定理得到，它指出：在连续映射下，弱收敛的性质是不变的.

定理 B.2.1(连续映射定理) 设$\{X,X_n\}$是取值于配备了 Borel-σ 代数的度量空间(S,d)上随机元，假定当 $n\to\infty$时，

$$X_n\Rightarrow X.$$

如果 φ：$S\to S'$是一个映射，且在 $X(\Omega)\subset S$ 上 a.s. 是连续的，其中，(S',d')是一个新的概率空间，那么，当 $n\to\infty$时，

$$\varphi(X_n)\xrightarrow{d}\varphi(X).$$

▶**例 B.2.2** 下面是一些重要的连续映射.

(ⅰ) 对某个 x_0，$\phi(f)=f(x_0)$，$f\in D([0,1];\mathbf{R})$.

(ⅱ) 对任意连续函数 g，$\phi(f)=\int g(f(s))\mathrm{d}s$，$f\in D([0,1];\mathbf{R})$.

(ⅲ) $\phi(f)=\sup_{t\in A} f(t)$，$\phi(f)=\sup_{t\in A}|f(t)|$，及相应的取下确界函数，其中 A 为一个紧集，$f\in C([0,1];\mathbf{R})$. 如果 $f\in D([0,1];\mathbf{R})$，那么，对满足 $f(a)=f(a-)$ 的 f，$\phi(f)=\sup_{t\leqslant a}|f(t)|$ 是连续的.

(ⅳ) 首次通过时间泛函 $\phi(f)=\inf\{t\geqslant 0: f(t)>a\}$，其中 a 为一个常数，$f\in\mathbf{C}([0,1];\mathbf{R})$. ◀

设 $\{S_1, d_1\}$，$\{S_2, d_2\}$ 是配备了 Borel-σ 代数的度量空间，定义乘积 $S_1\times S_2$ 为所有元素对 (f, g) 的集合，其中 $f\in S_1$，$g\in S_2$. 可以按照下列方式定义 $S_1\times S_2$ 上的度量：对 (f_1, g_1)，$(f_2, g_2)\in S_1\times S_2$，

$$d((f_1,g_1),(f_2,g_2)) = d_1(f_1,g_1)+d_2(f_2,g_2),$$

395

如果 S_1，S_2 都是可分的，那么，$S_1\times S_2$ 也是可分的.

乘积 σ 代数 $\mathcal{S}_1\otimes\mathcal{S}_2$ 是由所有(广义)矩形 $A_1\times A_2$ 生成的 σ 代数，其中，$A_1\in S_1$，$A_2\in S_2$. $\mathcal{B}(S_1\times S_2)$ 表示由所有开集 $O\subset S_1\times S_2$ 生成的 σ 代数. 如果 $\mathcal{S}_1\times\mathcal{S}_2$ 是可分的，那么，

$$\mathcal{B}(S_1\times S_2)=\mathcal{S}_1\otimes\mathcal{S}_2.$$

在以下的讨论中，所涉及的 σ 代数均为这个 σ 代数.

设 $\{X, X_n\}$，$\{Y, Y_n\}$ 是取值分别为 $\{S_1, d_1\}$，$\{S_2, d_2\}$ 的随机元，假定已经知道

$$X_n\Rightarrow X,\quad n\to\infty, \text{在}(S_1,d_1)\text{中}, \tag{B. 2}$$

$$Y_n\Rightarrow Y,\quad n\to\infty, \text{在}(S_2,d_2)\text{中}, \tag{B. 3}$$

下列定理给出了(B. 2)与(B. 3)成立的充分条件，这隐含了 (X_n, Y_n) 在乘积空间中联合弱收敛于 (X, Y).

定理 B. 2. 3(联合弱收敛) 设 $\{X, X_n\}$，$\{Y, Y_n\}$ 是取值分别为 $\{S_1, d_1\}$，$\{S_2, d_2\}$ 的两个序列，且(B. 2)，(B. 3)成立，那么，如果下列两个条件至少有一个成立：

(ⅰ) $Y=c\in S_2$ 是一个常数，即非随机的，

(ⅱ) 对所有 n，X_n 与 Y_n 是独立的，X 与 Y 也是独立的.

则必有

$$(X_n,Y_n)\Rightarrow(X,Y),\quad n\to\infty.$$

在上面的结果中取 $c=0$，结合连续映射定理，就得到下列度量空间中的 Slutzky 引理.

定理 B. 2. 4(SLUTZKY 引理) 设 X_n 有下列分解

$$X_n=Y_n+R_n,\quad n\geqslant 1,$$

其中，当 $n\to\infty$ 时，$R_n\Rightarrow 0$，或 $R_n\xrightarrow{P}0$，如果当 $n\to\infty$ 时，存在某个随机变量 Y，使得 $Y_n\Rightarrow Y$，则必有

$$X_n\Rightarrow Y,\quad n\to\infty.$$

下列结果经常被用到，但很少被提及.

396

定理 B. 2. 5 如果 $g:\mathbf{R}^{d_1}\to\mathbf{R}^{d_2}$ 是连续的，那么，映射 ϕ: $\mathcal{D}_\phi\to D([0,1];\mathbf{R}^{d_2})$ 关于 Skorohod 拓扑是连续的，其中 $\mathcal{D}_\phi\subset D([0,1];\mathbf{R}^{d_1})$ 由下列式子确定

$$\phi(f)(t)=g(f(t)),\quad f\in\mathcal{D}_\phi.$$

评注 B. 2. 6 在实际应用中，这个定理经常应用在：已经知道了当 $n\to\infty$ 时，$X_n\Rightarrow X$ 及 $P(X\in A)=1$，$A\subset D([0,1];\mathbf{R}^{d_1})$，$g$ 是定义在 A 上的连续映射，但是不一定属于 $D([0,1];\mathbf{R}^{d_1})$，那么可以通过设 $\mathcal{D}_\phi=A$ 来推出

$$\phi(X_n)\Rightarrow\phi(X),\quad n\to\infty.$$

特别地，可以推出 ϕ 除了一个 P_X-零集外是连续的.

下列表示定理是由 Skorohod，Dudley 和 Wichura 提出的，它建立了关于 ω 处处收敛的等价形式，它转变弱收敛为(几乎)处处依度量收敛，反之亦然. 它使得人们可以证明以概率 1 度量收敛来证明弱收敛. 读者可参考(Skorohod 和 Wellner，1986，p48，或 Billingsley，1999，定理 6. 7).

定理 B. 2. 7(Skorohod/Dudley/Wichura 表示定理)

(ⅰ) 设 (P_n, P) 是一列定义在度量空间 (S, d) 上的概率测度，满足 P 有一个可分的支集，或 (S, d)

是可分的，那么，存在定义在一个普通概率空间$(\widetilde{\Omega}, \widetilde{\mathcal{F}}, \mathbf{P})$上的随机元$(X_n, X)$，使得在$\mathbf{P}$下，对所有$n$，有$X \sim P$，$X_n \sim P_n$，且对所有$\omega \in \widetilde{\Omega}$，有

$$d(X_n(\omega), X(\omega)) \to 0, \quad n \to \infty.$$

(ⅱ) 设(X_n, X)是一列取值于度量空间(S, d)上的随机元，满足P_X有一个可分的支集，或(S, d)是可分的，那么，存在定义在一个普通概率空间$(\widetilde{\Omega}, \widetilde{\mathcal{F}}, \mathbf{P})$上的等价随机元$(\widetilde{X}_n, \widetilde{X})$，使得在$\mathbf{P}$下，对所有$n$，有$\widetilde{X} \sim P$，$\widetilde{X}_n \sim P_n$，且对所有$\omega \in \widetilde{\Omega}$，有

$$d(\widetilde{X}_n(\omega), \widetilde{X}(\omega)) \to 0, \quad n \to \infty.$$

(ⅲ) 逆命题：设(X_n, X)是一列取值于度量空间(S, d)上的随机元，P_X有一个可分的支集，或(S, d)是可分的，如对新的概率空间$(\widetilde{\Omega}, \widetilde{\mathcal{F}}, \mathbf{P})$上的等价随机元$\widetilde{X}_n \overset{d}{\Rightarrow} X_n$，$\widetilde{X} \overset{d}{\Rightarrow} X$，有

397

$$d(\widetilde{X}_n, \widetilde{X}) \overset{\mathbf{P}}{\to} 0, \quad n \to \infty,$$

那么，

$$X_n \Rightarrow X, \quad n \to \infty$$

成立.

B.3 Prohorov 定理

为了建立取值于空间$C([0, T]; \mathbf{R})$或空间$D([0, T]; \mathbf{R})$上的随机序列(X_n, X)的弱收敛，下面是一种基本方法：首先，建立有限维分布收敛，简记为 **fidi 收敛**，即对所有$k \in \mathbf{N}$，及所有的固定时间点$0 \leqslant t_1 < \cdots < t_k \leqslant T$，均有

$$(X_n(t_1), \cdots, X_n(t_k)) \overset{d}{\to} (X(t_1), \cdots, X(t_k)), \quad n \to \infty.$$

如果X是一个高斯过程，有限维分布收敛通常可用 Euclidean 空间$\mathbf{R}^k$中的中心极限定理来验证. fidi 收敛是建立弱收敛的一个必要条件，因为有限维分布决定了随机过程的分布，从而决定了该随机过程极限过程的分布. 不幸的是，fidi 收敛不是建立弱收敛的充分条件. 为了得到弱收敛的充分条件，必须讨论定义在$\mathcal{S}$上的概率测度空间$\mathcal{P}(S)$上的紧集的性质.

回忆一下，在度量空间(S, d)中，一个序列$\{x_n\}$被称为满足 Cauchy 性或是**基本的**，如果：$\forall \varepsilon > 0$，存在$n_0 = n_0(\varepsilon)$，使得当n，$m \geqslant n_0$时，有$d(x_n, x_n) < \varepsilon$成立. 如果一个度量空间$(S, d)$的所有基本列都是收敛的，则称该度量空间$(S, d)$是**完备的**. 每一个度量导出一个拓扑，可是两个不同的度量可能导出同一个拓扑，完备的度量空间有许多独有的性质. 如果一个度量空间为一个完备且可分的度量空间，则称该度量空间为 **Polish** 空间. 完备空间中紧集的性质可由闭集描述，因为在完备空间中任何闭子集是紧集的充要条件是该空间是**完全有界的**，即任何开球的覆盖均有有限子覆盖.

对一个配备了 Borel-σ代数$\mathcal{S}$及对应的概率测度集$\mathcal{P}(S)$的度量空间(S, d)，如$D([0, T]; \mathbf{R})$，$\mathcal{P}(\mathcal{S})$上的元素可用 **Prohorov 度量**π进行度量，关于 Prohorov 度量的收敛是弱收敛，即

$$P_n \Rightarrow P \quad 当且仅当 \quad \pi(P_n, P) \to 0.$$

在度量空间中，是否收敛可用序列判别法进行判别：一个序列收敛的充要条件是它的任一子列包含一个收敛的子列. 因此，概率测度序列(P_n, P)的弱收敛可由下列方法判定：当$n \to \infty$时，$P_n \Rightarrow P(\Leftrightarrow \pi(P_n, P) \to 0)$，当且仅当它的任一子列$\{P_{n_k}, k \geqslant 1\}$包含一个子列$\{P_{n'_k}, k \geqslant 1\}$，使得，当$k \to \infty$时，$P_{n'_k} \Rightarrow P(\Leftrightarrow \pi(P_{n'_k}, P) \to 0)$，显然$P_{n'_k} \Rightarrow P \Leftrightarrow P_{n_k} \Rightarrow P$，$k \to \infty$. 在度量空间中，如果子集$A$存在紧的闭包，则称$A$称为相对紧的. 如果$A$的任一子列$\{P_n\} \subset A$，均存在子列$\{P_{n_k}, k \geqslant 1\}$满足$P_{n_k} \Rightarrow P'$，$k \to \infty$，且极限$P'$属于$A$的闭包，则$A$是相对紧的. 将这个性质应用于概率测度空间，则得到：$A \subset \mathcal{P}(S)$有紧的闭包当且仅当

398

它的任一子列$\{P_n\} \subset A$，均存在子列$\{P_{n_k}, k \geqslant 1\}$弱收敛于某个$P' \in \overline{A}$，即$P_{n_k} \Rightarrow P'$，$k \to \infty$，其中，$P'$可

能依赖于子列. 因此，弱收敛 $P_n \Rightarrow P$，$n\to\infty$可以按下列方法证明：首先证明$\{P_n\}$是相对紧的，然后证明它的所有可能极限均为 P. 后者可以通过证明$\{P_n\}$依有限维分布收敛于 P 实现，因为$\{P_n\}$的有限维分布决定了极限 P 的分布. 下面的 Prohorov 定理建立了 S 中的紧集与$\mathcal{P}(S)$中紧集的关系，它是通过一个被称为**列紧**的概念实现的，$\mathcal{P}(S)$中的一个子集 $A\subset\mathcal{P}(S)$被称为是列紧的，如果对$\forall\varepsilon>0$，存在一个紧子集 $K_\varepsilon\subset S$，使得

$$P(K_\varepsilon) > 1-\varepsilon,\quad \text{对所有 } P\in A \text{ 成立}.$$

定理 B. 3. 1(Prohorov 定理) 设(S, d)是一个度量空间，则

(ⅰ) S 上任一概率测度的列紧集是相对紧的；

(ⅱ) 如果 S 是 *Polish* 空间，即完备且可分的度量空间，则任一相对紧集是列紧集.

注意，在 Prohorov 定理的第二个结论中，完备性是必需的. Prohorov 定理有下列推论.

推论 B. 3. 2 设$\{P_n\}$是度量空间(S, d)上列紧的概率测度序列，且$\{P_n\}$有限维分布收敛于 P，则

$$P_n\Rightarrow P,\quad n\to\infty.$$

B. 4 充分性准则

使用 Prohorov 定理，有限维分布收敛和列紧性可以证明弱收敛性.

取值于$C([0, 1];\mathbf{R})$上的过程 X，$X_n(n\geqslant 1)$，下列弱收敛的充分条件成立. 见(Billingsley，1968，定理 12. 3).

定理 B. 4. 1 设有限维分布收敛成立，即当 $n\to\infty$时，对所有 $k\in\mathbf{N}$，及所有时间点 t_1，…，$t_k\in[0, 1]$，均有

$$(X_n(t_1),\cdots,X_n(t_k))\xrightarrow{d}(X(t_1),\cdots,X(t_k)),\quad n\to\infty,\tag{B. 4}$$

且存在一个非减函数 F：$[0, 1]\to\mathbf{R}$，及两个常数 $\gamma\geqslant 0$，$\alpha>1$，使得对所有 s，$t\in[0, 1]$，均有

$$P(|X_n(t)-X_n(s)|\geqslant\lambda)\leqslant\lambda^{-\gamma}|F(t)-F(s)|^{\alpha},\tag{B. 5}$$

或

$$E(|X_n(t)-X_n(s)|^{\gamma})\leqslant|F(t)-F(s)|^{\alpha},\tag{B. 6}$$

对所有 $n\geqslant 1$ 和 $\lambda>1$ 成立，那么，当 $n\to\infty$时，有

$$X_n\Rightarrow X.$$

399

对取值于 $D([0, 1];\mathbf{R})$上的随机元序列$\{X, X_n\}$的弱收敛有一个类似的充分性准则，它需要下列准备知识：投影 π_t：$D([0, 1];\mathbf{R})\to\mathbf{R}$ 关于 x 是连续的，即如果 $d(x_n, x)\to 0$，$n\to\infty$，则有 $\pi_t(x_n)=x_n(t)\to x(t)=\pi_t(x)$，$n\to\infty$. 对连续投影，有下列结论：

$$\pi_t \text{ 关于 } x \text{ 是连续的当且仅当 } x \text{ 关于 } t \text{ 是连续的.}$$

事实上，如果 x 是连续的，则当 $d(x_n, x)\to 0$，$n\to\infty$时，设$\{\lambda_n\}$是满足$\|\lambda_n-\mathrm{id}\|_\infty$，$\|x_n-x\circ\lambda_n\|_\infty\to 0$，$n\to\infty$的变换，则有$|x_n(t)-x(t)|\leqslant|x_n(t)-x\circ\lambda_n(t)|+|x\circ\lambda_n(t)-x(t)|\to 0$，$n\to\infty$，其中，第二项收敛于 0 是由于 x 的连续性. 下面证明“仅当”部分，如果 x 关于 t 不连续，设 λ_n 是一个变换，它在$[0, t]$与$[t, 1]$上都是线性的，且满足 $\lambda_n(0)=0$，$\lambda_n(t)=t-1/n$ 和 $\lambda_n(1)=1$，那么，$x_n=x\circ\lambda_n$ 在 $D([0, 1];\mathbf{R})$中收于 x. 与 $x_n(t)$不收敛于 $x(t)$矛盾.

对一个取值于 $D([0, 1];\mathbf{R})$中的随机元 X，设

$$T_X=\{t\in[0,1]:\pi_t \text{ 关于 } x\in D([0,1];\mathbf{R})\setminus N \text{ 是连续的}\},$$

其中，$N\subset D([0, 1];\mathbf{R})$是一个 P_X-零集，即 $P(X\in\mathbf{N})=0$. 由右连左极函数在 0 点的右边续性知$0\in T_X$，但是，1 可能属于也可能不属于 T_X. 可以证明：存在最多可数个 t，使得

$$P(X\text{ 在 }t\text{ 点不连续}) = P(X \in J_t) > 0,$$

其中，J_t 表示在 t 点不连续的右连左极函数. 见(Billingsley，1999，p. 138).

对 t_1，…，$t_k \in T_X$，则投影

$$\pi_{t_1,\cdots,t_k}(X) = (X(t_1),\cdots,X(t_k))$$

在满足 $A \subset D([0,1];\mathbf{R})$，$P(X \in A)=1$ 的集合 A 上是连续的. 下列充分性准则可在(Billingsley，1968，定理 15.6)或(Billingsley，1999，定理 13.5)中找到.

定理 B.4.2 假设：

(ⅰ) 在下列意义下，有限维分布收敛在连续点成立：当 $n\to\infty$ 时，对所有 $k\in\mathbf{N}$，及所有 t_1，…，$t_k \in T_X$，均有

$$(X_n(t_1),\cdots,X_n(t_k)) \Rightarrow (X(t_1),\cdots,X(t_k)). \tag{B.7}$$

(ⅱ) 当 $\delta\to 0$ 时，

$$X_1 - X_{1-\delta} \Rightarrow 0, \tag{B.8}$$

或

$$P(X_1 \neq X_{1-}) = 0.$$

(ⅲ) 对所有 $r\leqslant s\leqslant t$ 和 $\lambda>0$ 有

400

$$P(|X_n(s)-X_n(t)| \geqslant \lambda, |X_n(t)-X_n(s)| \geqslant \lambda) \leqslant \frac{1}{\lambda^{4\beta}}[F(t)-F(r)]^{2\alpha}, \tag{B.9}$$

或

$$E|X_n(s)-X_n(t)|^{\beta}|X_n(t)-X_n(s)|^{\beta} \leqslant [F(t)-F(r)]^{2\alpha}, \tag{B.10}$$

其中，$\beta\geqslant 0$，$\alpha>1/2$，F 为$[0,1]$上非减的连续函数.

那么，当 $n\to\infty$ 时，有

$$X_n \Rightarrow X.$$

注意，(B.8)式等价于

$$\lim_{\delta\downarrow 0}\limsup_{n\to\infty} P(|X_n(1-\delta)-X_n(1)| > \varepsilon) = 0 \tag{B.11}$$

对所有 $\varepsilon>0$ 成立.

对路径在 Skorohod 空间 $D([0,\infty);\mathbf{R}^d)$中的过程，有下列弱收敛定理，见(Jacod 和 Shiryaev，2003，P. 319).

定理 B.4.3 除了假定有限维分布收敛在连续点成立外，还假定(B.11)式成立，且存在$(0,\infty)$上非减的函数 F 和常数 $\gamma\geqslant 0$，$\alpha>1$ 使得对任意 $\lambda>0$ 和 $r<s<t$ 均有

$$P(\|X_n(r)-X_n(s)\| \geqslant \lambda, \|X_n(t)-X_n(t)\| \geqslant \lambda) \leqslant \lambda^{-\gamma}[F(t)-F(r)]^{\alpha} \tag{B.12}$$

对所有 $n\geqslant 1$ 成立. 那么，当 $n\to\infty$ 时，

$$X_n \Rightarrow X$$

在 $D([0,\infty);\mathbf{R}^d)$中成立.

引理 B.4.4 设取值于 $D([0,\infty);\mathbf{R}^d)$中的过程 X，X_n，$n\geqslant 1$ 有独立增量，且

(ⅰ) 当 $n\to\infty$ 时，对所有 $t>0$，有 $X_n(t)\xrightarrow{d}X(t)$；

(ⅱ) 当 $n\to\infty$ 时，对所有 $s<t$，有 $X_n(t)-X_n(s)\xrightarrow{d}X(t)-X(s)$，

则有限维分布收敛成立.

B.5 Skorohod 空间的进一步讨论

前面定义的 Skorohod 度量，也称为 J_1 度量，满足跳收敛性质. 事实上，如果当 $n\to\infty$ 时，在 J_1 拓

扑下有 $x_n\to x$，那么，对任意内点 t，存在点列$\{t_n\}$，使得当 $n\to\infty$时，$t_n\to t$，$x_n(t_n)\to x(t)=x(t+)$，且 $x_n(t_n-)\to x(t-)$，这意味着，当 $n\to\infty$时，$x_n(t_n)-x_n(t_n-)\to x(t)-x(t-)$，即满足跳收敛性质. 从而，当 $t<1/2-1$ 时，函数 $f_n(t)$为 0，当 $t>1/2$ 时，函数 $f_n(t)=1$，且为线性插值函数，否则，在 J_1 拓扑下不收敛于示性函数 $1_{[1/2,1]}(t)$，有时，这种度量的条件太强了.

对于不要求跳收敛的情形，M_1 度量是一个不错的选择. 设 $f\in D([0,1];\mathbf{R})$，按下列方式定义完全图

401

$$\Gamma_f=\{(z,t)\in\mathbf{R}\times[0,1]:z=\alpha f(t-)+(1-\alpha)f(t),\text{对某个 }\alpha\in[0,1]\}.$$

其中用到了默认条件 $f(0-)=f(t)$. 注意到：Γ_f 是一个连通集，且对所有跳点 t，Γ_f 包含了连接$(f(t-),t)$与$(f(t),t)$之间的所有线性连线. 可以按下列方式定义 Γ_f 的(总)序，如果 $t_1<t_2$，或 $t_1=t_2$，且 $|x(t_1-)-z_1|\leqslant|x(t_2-)-z_2|$，则$(z_1,t_1)\leqslant(z_1,t_1)$. 按照这种排序方法，该完全图始于左边，终于右边. Γ_f 的一个参数表示是一个非减连续函数：θ：$[0,1]\to\Gamma_f$. 记 $\theta=(u,t)$，即对 $s\in[0,1]$，$\theta(s)=(u(s),t(s))$，其中，$u(s)$表示空间坐标，$t(s)$表示时间坐标. 用 Π_f 表示 Γ_f 的所有参数表示的集合，则 M_1 度量定义如下

$$d_{M_1}(f_1,f_2)=\inf\max(\|u_1-u_2\|_\infty,\|t_1-t_2\|_\infty),$$

其中，下确界是对所有$(u_1,t_1)\in\Pi_{f_1}$，及$(u_2,t_2)\in\Pi_{f_2}$取. 如果 f 是连续函数，且当 $n\to\infty$时，$d_{M_1}(f_n,f)\to 0$，则同样有$\|f_n-f\|_\infty\to0$，$n\to\infty$. J_1 拓扑比 M_1 拓扑强同，即在 J_1 拓扑下收敛必定在 M_1 拓扑下收敛量.

一个更弱的收敛是 M_2 度量下的收敛，它建立在完全图之间的 Hausdorff 距离之上. 设 K_1，K_2 是两个紧集，**Hausdorff 度量**定义为

$$d_H(K_1,K_2)=\max\{\sup_{f\in K_1}d(f,K_2),\sup_{f\in K_2}d(f,K_1)\}.$$

其中，$d(x,A)=\inf\{d(x,y):y\in A\}$表示点 x 与集合 A 之间的距离. M_2 度量定义如下

$$d_{M_2}(f,g)=d_H(\Gamma_f,\Gamma_g).$$

显然在 J_1 拓扑或 M_1 拓扑下收敛必定在 M_1 拓扑下收敛.

B.6 鞅差分的中心极限定理

本节讨论定理 8.3.6 的替换与推广，即鞅差分的中心极限定理. 下列结果可以在 Durrett(1996)中找到.

定理 B.6.1(鞅差阵列的 Lindberg-Feller CLT) 设$\{X_{nk}:1\leqslant k\leqslant n,n\geqslant1\}$是 F_{nk}-鞅差阵列，满足对所有 n，k 均有 $E(X_{nk}^2|F_{nm})<\infty$，记 $V_{nk}=\sum_{i=1}^{k}E(X_{ni}^2|F_{n,i-1})$，$S_n=\sum_{k=1}^{n}X_{nk}$. 如果下列两个条件成立，

(i) $V_{nn}\to1$(依率)，

(ii) Lindberg 条件成立，即对所有 $\varepsilon>0$ 依概率有

$$\sum_{k=1}^{n}E(X_{nk}^2\mathbf{1}(|X_{nk}|>\varepsilon)|\mathcal{F}_{n,k-1})\to0,$$

那么，当 $n\to\infty$时，有 $S_n\xrightarrow{d}N(0,1)$.

402

对鞅差序列，下列结果是直接的.

定理 B.6.2(鞅差序列的 Lindberg-Feller CLT) 设$\{X_n\}$是一个$\mathcal{F}_n$-鞅序列，记 $V_n=\sum_{k=1}^{n}E(X_k^2|\mathcal{F}_{k-1})$，$S_n=\sum_{k=1}^{n}X_k$，如果

(i) $V_T/T\to\sigma^2>0$；

（ⅱ）Lindberg 条件成立，即当 $n\to\infty$ 时，对任意 $\varepsilon>0$，有

$$\frac{1}{n}\sum_{k=1}^{n}E(X_k^2\mathbf{1}(|X_k|>\varepsilon))\to 0,$$

那么，当 $n\to\infty$ 时，有 $S_n/\sqrt{n}\xrightarrow{d}N(0,\sigma^2)$.

评注 B.6.3 对应的泛函中心极限定理也是成立的. 特别地，在定理 B.7.2 中将指出，在没有改变条件的情况下，结论将增强为：当 $n\to\infty$ 时，有 $S_{\lfloor nt\rfloor}/\sqrt{n}\Rightarrow B$.

定理 B.6.4(鞅差阵列的 Lindberg-Feller FCLT) 将定理 B.7.2 的条件(ⅰ)推广为：对 $t\in[0,1]$，$V_{n,\lfloor nt\rfloor}\to t$(依概率)，那么，部分和过程弱收敛于布朗运动即，当 $n\to\infty$ 时，$S_{n,\lfloor nt\rfloor}\Rightarrow B(t)$.

B.7 泛函中心极限定理

泛函中心极限定理，也称不变原理，是对来自金融市场的时间序列进行统计推断的核心数学定理. 泛函中心极限定理也是进行估计和推断这种统计过程中的基本工具，因为在这个过程中的许多统计量，特别是部分和过程，都是由一些基本随机过程驱动的，它们满足某个泛函中心极限定理，它限定统计量的极限是某个有趣的分布.

Donsker 的经典不变原理在部分序列与在第 4 章的结尾已经讨论过的布朗运动之间提供了一个基本关系. 它适用于增量是均值为 0，且有有限的二阶矩的独立同分布的随机变量的情形.

定理 B.7.1(Donsker，I. I. D. 情形) 设 ξ_1，ξ_2，…是一列 i. i. d. 随机变量，且 $E(\xi_1)=0$，$\sigma^2=E(\xi_1^2)<\infty$，那么，当 $T\to\infty$ 时，

$$\frac{1}{\sqrt{T}}\sum_{t=1}^{\lfloor T\cdot\rfloor}\xi_t\Rightarrow\sigma B(\cdot),$$

403

其中，B 表示标准布朗运动，"$\Rightarrow$"表示在 Skorohod 空间 $D([0,1];\mathbf{R})$ 中弱收敛.

Donsker 定理适用于独立同分布的随机变量，在金融问题中这个条件太强了，为了研究数理金融及金融市场中的问题，将其推广到相依随机变量的情形是必需的. 下面是 Lindberg 条件下，一个鞅差序列的对应结果.

定理 B.7.2(鞅差序列的 Lindberg-Feller FCLT) 设 $\{\xi_t\}$ 是一个平方可积的 $\mathcal{F}_t$-鞅差序列，$V_t=\sum_{k=1}^{t}E(\xi_k^2\,|\,\mathcal{F}_{k-1})$，如果

（ⅰ）当 $T\to\infty$ 时 $V_T/T\to\sigma^2>0$(依概率)，

（ⅱ）Lindberg 条件成立，即当 $T\to\infty$ 时，对任意 $\varepsilon>0$，有

$$\frac{1}{T}\sum_{t=1}^{T}E(\xi_t^2\mathbf{1}(|\xi_t|>\sqrt{T}\varepsilon))\to 0,$$

那么，当 $T\to\infty$ 时，有

$$\frac{1}{\sqrt{T}}\sum_{t=1}^{\lfloor T\cdot\rfloor}\xi_t\Rightarrow\sigma B(\cdot).$$

下面是推广到鞅差阵列的情形.

定理 B.7.3(鞅差阵列的 Lindberg-Feller FCLT) 设 $\{\xi_{Tt}:1\leqslant t\leqslant T,\ T\geqslant 1\}$ 是 $\mathcal{F}_{Tt}$-鞅差阵列，且对所有 $1\leqslant t\leqslant T$，$T\geqslant 1$，均有 $E(\xi_{Tt}^2\,|\,\mathcal{F}_{T,t-1})<\infty$，记

$$V_{tk}=\sum_{i=1}^{k}E(\xi_{ti}^2\,|\,\mathcal{F}_{t,i-1}),\quad 1\leqslant t\leqslant T,T\geqslant 1.$$

如果下列两个条件成立，

（ⅰ）对所有 $u\in[0,1]$，均有 $V_{T,\lfloor Tu\rfloor}\to u$(依概率)，

（ⅱ）Lindberg 条件成立，即当 $T\to\infty$ 时，依概率有

$$\sum_{i=1}^{T}E(\xi_{Ti}^2\mathbf{1}(|\xi_{Ti}|>\varepsilon)|\mathcal{F}_{T,i-1})\to 0,$$

对所有 $\varepsilon>0$ 成立.

那么，当 $T\to\infty$ 时，有

$$\sum_{t=1}^{\lfloor Tu\rfloor}\xi_{Tt}\Rightarrow B(u).$$

注意，定理 $B.7.2$ 是定理 B.7.3 的特殊情形.

404

将 α 混合序列的中心极限定理推广到泛函情形，就得到下列结果，见 Ibragimov(1962)及(Hall 和 Heyde，1980，推论 5.1).

定理 B.7.4（α 混合序列的 FCLT） 设 $\{X_n: n\in\mathbf{Z}\}$ 是一个遍历平稳的序列，满足 $E(X_0)=0$，且存在 $\delta>0$，使得 $E|X_0|^{2+\delta}<\infty$. 又设 α-混合系数 $\alpha(k)$，$k\in\mathbf{N}$ 满足

$$\sum_{k=1}^{\infty}\alpha(k)^{\frac{\delta}{2(2+\delta)}}<\infty.$$

那么，当 $T\to\infty$ 时，有

$$\frac{1}{\sqrt{T}}\sum_{t=1}^{\lfloor Ts\rfloor}X_t\Rightarrow\sigma B_s,$$

其中，

$$\sigma^2=E(X_0^2)+2\sum_{k=1}^{\infty}E(X_1X_{1+k}).$$

下列关于长记忆过程的结果来自 Taqqu(1974/75).

定理 B.7.5（长记忆过程的 FCLT） 设 $\{X_t: t\in\mathbf{N}\}$ 是一个高斯随机变量构成的平稳序列，且其均值为 0，自协方差函数为 $\gamma(k)=E(X_1X_{1+k})$，$k\in\mathbf{Z}$. 假定

$$\lim_{T\to\infty}\frac{1}{KT^{2H}L(T)}\sum_{t=1}^{T}\sum_{s=1}^{T}\gamma(t-s)=1,$$

对某个 $0<H<1$，某个常数 $K>0$ 及某个在无穷远处缓变的函数 L 成立，那么，

$$\frac{1}{\sqrt{KT^{2H}L(T)}}\sum_{t=1}^{\lfloor Ts\rfloor}X_t\Rightarrow B_s^H,$$

其中，$\{B_t^H: t\in[0,1]\}$ 是一个标准分数布朗运动.

B.8 强逼近

设 ξ_t，因此 S_T，定义在普通概率空间$(\Omega,\mathcal{F},P)$上，由 S_T 弱收敛于布朗运动 B 受到启发，是否可以在$(\Omega,\mathcal{F},P)$上定义一个布朗运动，使 S_T 在$[0,1]$上是一致收敛于 B 的呢？与之相关的问题是部分和 $\sum_{t=1}^{n}\xi_t$ 是否逼近 $B(n)$. 如果有这样的逼近，是否足够地强，足以保证 Donsker 定理意义下的不变原理成立呢？这些问题在经典概率理论中已经被广泛地研究过了.

405

下列定理为部分和逼近布朗运动提供了一些众所周知的结果，这些结果也被称为**强不变原理**，它们具有下列典型的形式

$$E_n=o(a_n),\quad \text{a.s.},$$

其中，E_n 是一个包含布朗运动的随机表达式. 上式意味着

$$P(\lim_{n\to\infty} E_n/a_n = 0) = 1.$$

定理 B. 8. 1(强逼近，I. I. D. 情形) 设 ξ_1，ξ_2，…是一列 i. i. d. 随机变量，且 $E(\xi_1)=0$，$\sigma^2=E(\xi_1^2)<\infty$，设 $S(n)=\sum_{i=1}^{n}\xi_i$，$n\in\mathbf{N}$，则

(ⅰ) 可以定义一个标准布朗运动$\{B(t),\ 0\leqslant t<\infty\}$，使得

$$\max_{n\leqslant N}|S(n)-B(n)| = o((N\log\log N)^{1/2}), \quad \text{a. s.}$$

(ⅱ) 如果存在 $r>2$，使得 $E|\xi_1|^r<\infty$，那么，

$$|S(n)-B(n)| = o(n^{1/r}), \quad \text{a. s.}$$

(ⅲ) 如果 H：$[0,\infty)\to[0,\infty)$是一个函数，满足 $t^{-2}H(t)$是非减的，$t^{-3}H(t)$是非增的，且 $EH(\xi_1)<\infty$，那么，

$$\left|\sum_{i=1}^{n}\xi_i - B\left(\sum_{i=1}^{n}\sigma_i^2\right)\right| = o(H^{-1}(n)),$$

其中，满足 $\sigma_n\to 1$ 的实数序列$\{\sigma_n:\ n\in\mathbf{N}\}$可以按下列方式选择

$$1-\sigma_n^2 = o(n^{-1}(H^{-1}(n))^2).$$

由上面部分和的强逼近可以推出

$$N^{-1/2}\max_{n\leqslant N}|S(n)-B(n)| \xrightarrow{P} 0, \quad (N\to\infty), \tag{B. 13}$$

通常称之为**弱不变原理**. 但它也是足够地强，足以保证 Donsker 定理意义下(即弱收敛意义下)的不变原理成立.

定理 B. 8. 2 设 ξ_1，ξ_2，…满足(B. 13)意义下的不变原理，那么，当 $T\to\infty$时，

$$S_T(u) = T^{-1/2}S(\lfloor Tu\rfloor)\Rightarrow B(u).$$

406

证明 证明将用到 Skorohod/Dudley/Wichura 表示定理. 由比例性质(5. 2. 10)知：对每个 $T>0$，$\{B(u):u\geqslant 0\}$与$\{T^{-1/2}B(Tu):\ u\geqslant 0\}$的分布相同，因此，

$$\sup_{u\in[0,1]}|S_T(u)-B(u)| \overset{d}{=} \sup_{u\in[0,1]}\left|\frac{1}{\sqrt{T}}\sum_{t=1}^{\lfloor Tu\rfloor}\xi_t - \frac{1}{\sqrt{T}}B(Tu)\right|. \tag{B. 14}$$

所以，当 $T\to\infty$时，在一个新的概率空间上有

$$\begin{aligned}\sup_{u\in[0,1]}|S_T(u)-B(u)| &= \sup_{u\in[0,1]}\frac{1}{\sqrt{T}}\left|\sum_{t=1}^{\lfloor Tu\rfloor}\xi_t - B(Tu)\right| \\ &= \frac{1}{\sqrt{T}}\max_{n\leqslant T}|S(n)-B(n)| \\ &\xrightarrow{P} 0,\end{aligned}$$

■

由此得到：当 $T\to\infty$时，对原始过程有 $S_T\Rightarrow B$.

参考文献

Billingsley P. (1968) *Convergence of Probability Measures*. John Wiley & Sons Inc., New York.

Billingsley P. (1999) *Convergence of Probability Measures*. Wiley Series in Probability and Statistics: Probability and Statistics 2nd edn. John Wiley & Sons Inc., New York. A Wiley-Interscience Publication.

Durrett R. (1996) *Probability: Theory and Examples*. 2nd edn. Duxbury Press, Belmont, CA.

Hall P and Heyde C.C. (1980) *Martingale Limit Theory and its Application*. Academic Press Inc. [Harcourt Brace Jovanovich Publishers], New York. Probability and Mathematical Statistics.

Ibragimov I.A. (1962) Some limit theorems for stationary processes. *Teor. Verojatnost. i Primenen.* **7**, 361–392.

Jacod J. and Shiryaev A.N. (2003) *Limit Theorems for Stochastic Processes*. vol. 288 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]* 2nd edn. Springer-Verlag, Berlin.

Pollard D. (1984) *Convergence of Stochastic Processes*. Springer Series in Statistics. Springer-Verlag, New York.

Shiryaev A.N. (1999) *Essentials of Stochastic Finance: Facts, Models, Theory*. vol. 3 of *Advanced Series on Statistical Science & Applied Probability*. World Scientific Publishing Co. Inc., River Edge, NJ. Translated from the Russian manuscript by N. Kruzhilin.

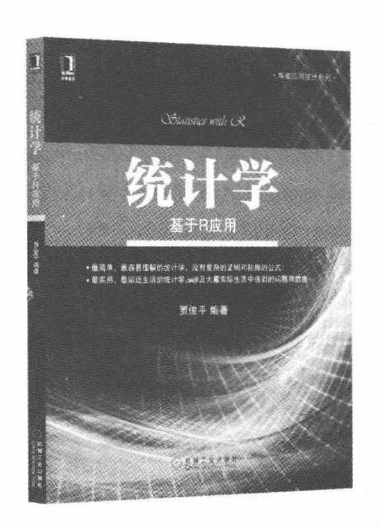
Shorack G.R. and Wellner J.A. (1986) *Empirical Processes with Applications to Statistics*. Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics. John Wiley & Sons Inc., New York.

Taqqu M.S. (1974/75) Weak convergence to fractional Brownian motion and to the Rosenblatt process. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* **31**, 287–302.

407

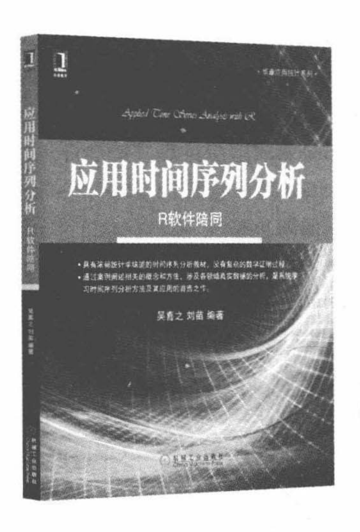
索　引

推荐阅读



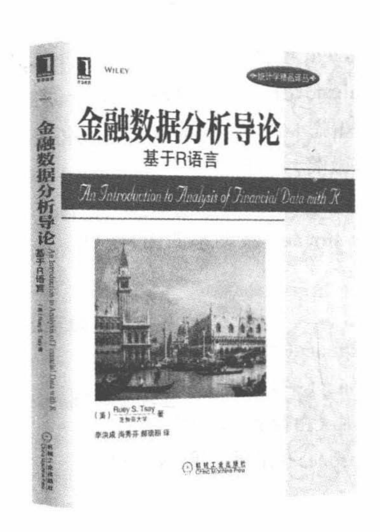
统计学：基于R应用

作者：贾俊平 ISBN：978-7-111-46651-2 定价：39.00元



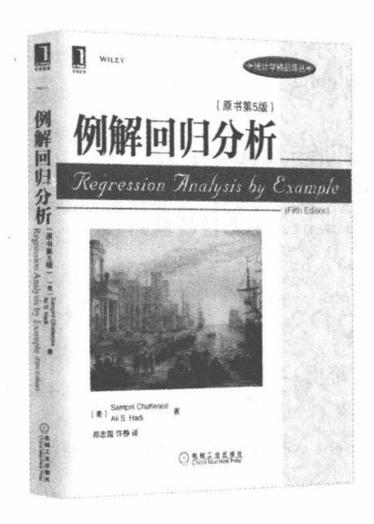
应用时间序列分析：R软件陪同

作者：吴喜之 刘苗 ISBN：978-7-111-46816-5 定价：39.00元



金融数据分析导论：基于R语言

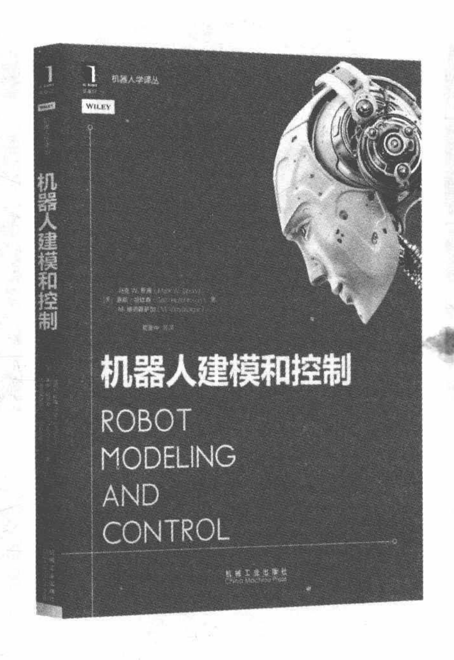
作者：Ruey S. Tsay ISBN：978-7-111-43506-8 定价：69.00元



例解回归分析（原书第5版）

作者：Samprit Chatterijee 等 ISBN：978-7-111-43156-5 定价：69.00元

推荐阅读



机器人建模和控制

书号：978-7-111-54275-9 作者：马克 W. 斯庞 等 译者：贾振中 等 定价：79.00元

本书所覆盖的内容在深度和广度方面都是独一无二的。据我所知，没有其他教材能够对现代机器人的操作和控制做出如此精彩而又全面的概述。

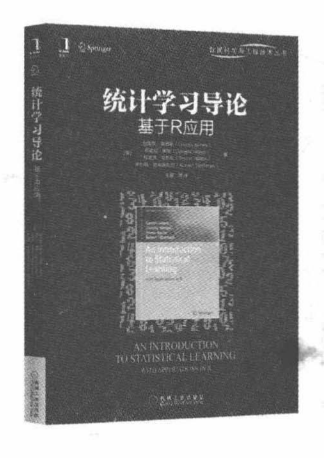
——布拉德利·毕晓普（Bradley Bishop），美国海军学院

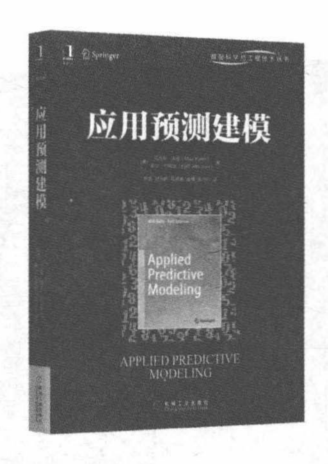
本书由Mark W. Spong、Seth Hutchinson和M. Vidyasagar三位机器人领域顶级专家联合编写，全面且深入地讲解了机器人的控制和力学原理。全书结构合理、推理严谨、语言精练，习题丰富，已被国外很多名校（包括伊利诺伊大学、约翰霍普金斯大学、密歇根大学、卡内基-梅隆大学、华盛顿大学、西北大学等）选作机器人方向的教材。

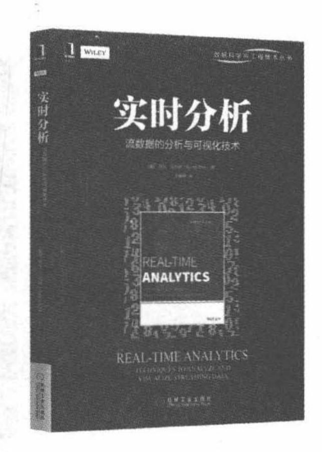
本书特色

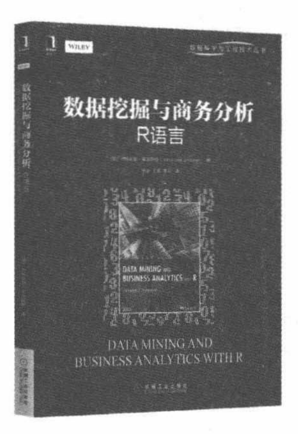
- 通过循序渐进的计算方法帮助你推导和计算最通用机器人设计中的运动学正解、运动学逆解和雅克比矩阵问题。
- 详细覆盖了计算机视觉和视觉伺服控制，使你能够通过对带有相机感知元件的机器人编程来操作物体。
- 通过一个完整章节的动力学讲解，为计算最通用机械臂设计中的动力学问题做好准备。
- 初步介绍了最通用的运动规划和轨迹。

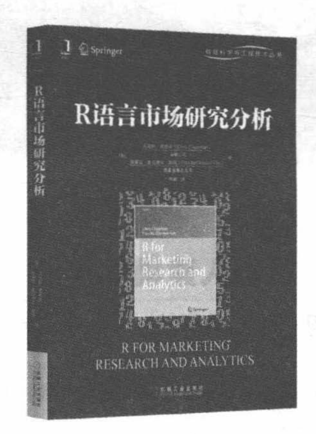
推荐阅读

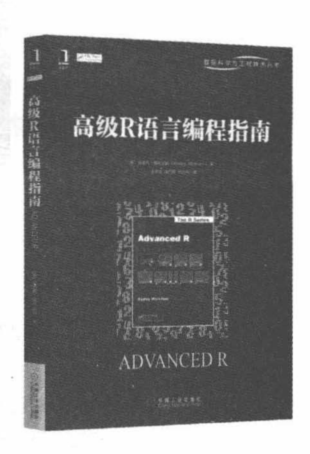












统计学习导论——基于R应用

作者：加雷斯·詹姆斯 等 ISBN：978-7-111-49771-4 定价：79.00元

应用预测建模

作者：马克斯·库恩 等 ISBN：978-7-111-53342-9 定价：99.00元

实时分析：流数据的分析与可视化技术

作者：拜伦·埃利斯 ISBN：978-7-111-53216-3 定价：79.00元

数据挖掘与商务分析：R语言

作者：约翰尼斯·莱道尔特 ISBN：978-7-111-54940-6 定价：69.00元

R语言市场研究分析

作者：克里斯·查普曼 等 ISBN：978-7-111-54990-1 定价：89.00元

高级R语言编程指南

作者：哈德利·威克汉姆 ISBN：978-7-111-54067-0 定价：79.00元

金融统计与数理金融 方法、模型及应用

数理金融已经发展成为一个庞大的研究领域，同时它也需要大量复杂的数学工具。本书系统而深入地介绍了各种金融方法和相关的数学工具，论证严谨，案例翔实，适合高水平的研究人员以及对数理金融或金融统计感兴趣的实践工作者阅读，也适合统计、金融、计量经济学和工商管理等专业的本科生和研究生阅读。

本书特点

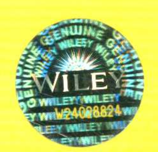
- 介绍金融统计和金融数学的基础知识。
- 解释统计方法在计量经济学和金融工程中的重要性及用途。
- 强调导数和微积分的作用以帮助理解方法和结果。
- 涉及较深的理论，如鞅论、随机过程、随机积分。
- 通过丰富的实例说明数学和统计学在金融中的应用。
- 配套网站提供与本书内容相关的R代码和数据集。

作者简介

安斯加尔·斯特兰（Ansgar Steland）德国数理经济学会、计量经济学会、生物统计学会、社会政治联盟等学会的会士，现为德国名校亚琛工业大学的教授，研究领域包括时间序列分析、数理经济学、统计计算、应用数理统计和金融统计等。Steland于1996年从德国哥廷根大学博士毕业，师从Manfred Denker。Steland在学术上非常活跃，已发表几十篇优秀学术论文，并被广泛引用。曾应邀到世界各地做学术报告，包括美国斯坦福大学、奥地利因斯布鲁克大学、荷兰马斯特里赫特大学、捷克布拉格查理大学、德国哥廷根大学等。

Financial Statistics and Mathematical Finance

Methods, Models and Applications



WILEY
www.wiley.com

上架指导：数学

ISBN 978-7-111-57301-2

9 787111 573012

定价：85.00元

投稿热线：(010) 88379604
客服热线：(010) 88378991 88361066
购书热线：(010) 68326294 88379649 68995259

华章网站：www.hzbook.com
网上购书：www.china-pub.com
数字阅读：www.hzmedia.com.cn

封面设计·杨宇梅